

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 25, 2003, 12:10:38 ; Search time 2162 Seconds  
(without alignments)  
8969.086 Million cell updates/sec

Title: US-10-049-568-1

Perfect score: 474

Sequence: 1 gccagattatttcagtggc.....ggagatgcacctgagttaa 474

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb\_ba.\*

2: gb\_hg.\*

3: gb\_in.\*

4: gb\_om.\*

5: gb\_ov.\*

6: gb\_pat.\*

7: gb\_ph.\*

8: gb\_pl.\*

9: gb\_pr.\*

10: gb\_ro.\*

11: gb\_sts.\*

12: gb\_sy.\*

13: gb\_un.\*

14: gb\_vi.\*

15: em\_ba.\*

16: em\_fun.\*

17: em\_hum.\*

18: em\_in.\*

19: em\_mu.\*

20: em\_om.\*

21: em\_or.\*

22: em\_ov.\*

23: em\_pat.\*

24: em\_ph.\*

25: em\_pl.\*

26: em\_ro.\*

27: em\_sts.\*

28: em\_un.\*

29: em\_vi.\*

30: em\_hig\_hum.\*

31: em\_hig\_inv.\*

32: em\_hig\_other.\*

33: em\_hig\_mus.\*

34: em\_hig\_pln.\*

35: em\_hig\_rod.\*

36: em\_hig\_nam.\*

37: em\_hig\_vrt.\*

38: em\_sy.\*

39: em\_higo\_hum.\*

40: em\_higo\_mus.\*

41: em\_higo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES.

Result No.	Score	Query Match %	Length	DB	ID	Description
1	474	100.0	474	6	AX088165	Sequence
2	439.2	92.7	1191	6	AX128507	Sequence
3	439.2	92.7	2274	6	AX549337	Sequence
4	439.2	92.7	2274	9	AF190500	Homo sapi
5	439.2	92.7	2467	6	BD135245	Novel mam
6	439.2	92.7	3584	6	BD135246	Novel mam
7	222	46.8	187431	9	AC107219	Homo sapi
8	171.8	36.2	1665	6	BD141178	Novel G p
9	171.8	36.2	1668	6	AX148176	Sequence
10	171.8	36.2	1473	6	BD141176	Novel G p
11	171.8	36.2	1545	6	BD141175	Novel G p
12	171.8	36.2	1830	6	BD141172	Novel G p
13	171.8	36.2	2049	6	AX385037	Sequence
14	171.8	36.2	2142	6	AX451566	Sequence
15	171.8	36.2	2181	6	AX709190	Sequence
16	171.8	36.2	2190	6	BD141181	Novel G p
17	171.8	36.2	2193	6	AX385032	Sequence
18	171.8	36.2	2214	6	AX451562	Sequence
19	171.8	36.2	2262	6	BD141180	Novel G p
20	171.8	36.2	2265	6	AX385027	Sequence
21	171.8	36.2	2436	9	AF453828	Homo sapi
22	171.8	36.2	2838	9	AF403384	Homo sapi
23	170.6	36.0	220994	2	AC128964	Rattus no
24	170.6	36.0	230365	2	AC121415	Rattus no
25	170.6	36.0	243409	2	AC098607	Rattus no
26	163	34.4	158394	2	AC122462	Mus muscu
27	163	34.4	253348	2	AC132143	Mus muscu
28	156.8	33.1	321	6	AX384516	Sequence
29	146.2	30.8	2214	6	AX385045	Sequence
30	146.2	30.8	2539	10	AF346501	Sequence
31	106.8	22.5	1018	6	AX147820	Sequence
32	106.8	22.5	1018	6	AX521869	Sequence
33	106.8	22.5	44585	6	AX646897	Sequence
34	106.8	22.5	170522	9	AL138708	Human DNA
35	92.4	19.5	180673	10	AC068627	Mus Muscu
36	92.4	19.5	213462	10	AC077689	Mus Muscu
37	86	18.1	211251	2	AC121029	Rattus no
38	86	18.1	239190	2	AC098990	Rattus no
39	73	15.4	167899	2	BX005476	Danio rer
40	73	15.4	195380	5	AL923075	Zebrafish
41	61.4	13.0	176565	2	AL136106	Homo sapi
42	57.2	12.1	653	4	AY196483	Equus cab
43	48.8	10.3	1015	6	AX254453	Sequence
44	47	9.9	104992	2	AC005504	Plasmodiu
45	47	9.9	169546	2	AC004157	Plasmodiu

ALIGNMENTS

RESULT 1	AX088165	AX088165	Sequence 1 from Patent WO0114548.	474 bp	DNA	linear	PAT 17-MAR-2001
LOCUS	AX088165	Sequence 1 from Patent WO0114548.					
DEFINITION	AX088165	Sequence 1 from Patent WO0114548.					
ACCESSION	AX088165	Sequence 1 from Patent WO0114548.					
VERSION	AX088165.1	GI:13397078					
KEYWORDS							
SOURCE		Homo sapiens (human)					
ORGANISM		Homo sapiens					
		Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;					
		Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
REFERENCE							
AUTHORS		Duecker K.					
TITLE		New g-protein coupled receptor and dna sequences thereof					
JOURNAL		Patent: WO 0114548-A 1 01-MAR-2001;					

FEATURES  
source  
Location/Qualifiers  
1..474  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
<1..474

CDS  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAC34686.1"  
/db\_xref="GI:13397079"  
/translation="AQIYSAIFGILNLAFLIIVFSYGMFYSVHOSATIAPEIRNQ  
VKEMILAKRFIFIFDALCWFIFVAKPLSLQVEIPGITSWVYIGYSAINSLN  
PILYITTPFKEMIHREWHNYRQKSMDSKIRKMLHSHSGKCGHCRCHLS"  
BASE COUNT 144 a 92 c 88 g 150 t  
ORIGIN

Query Match 100.0%; Score 474; DB 6; Length 474;  
Best Local Similarity 100.0%; Pred. No. 1e-99;  
Matches 474; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GCCAGATTATTCAGTGGCAATTTCTTGGTATTAAATTTGGCGCATTTATCATATA 60  
Db 1 GCCAGATTATTCAGTGGCAATTTCTTGGTATTAAATTTGGCGCATTTATCATATA 60  
QY 61 GTTTTCTCTATGGAAGCATGTTTATATAGTGTTCATCAAAAGTCCATACAGCAACTGAA 120  
Db 61 GTTTTCTCTATGGAAGCATGTTTATATAGTGTTCATCAAAAGTCCATACAGCAACTGAA 120  
QY 121 ATACCGAATCAAGTAAAAAGAGATGATCTTCCCAAGCTTTTCTTATATAGTATT 180  
Db 121 ATACCGAATCAAGTAAAAAGAGATGATCTTCCCAAGCTTTTCTTATATAGTATT 180  
QY 181 ACTGATGATTATCTGGTATACCATTTTGTAGCGAAACCTCTTTCACTGCTTCAGGTA 240  
Db 181 ACTGATGATTATCTGGTATACCATTTTGTAGCGAAACCTCTTTCACTGCTTCAGGTA 240  
QY 241 GAAATACAGGTACCAATACCTCTTGGGTAGTATGTTGTTATTCGCCATTAACAGTCT 300  
Db 241 GAAATACAGGTACCAATACCTCTTGGGTAGTATGTTGTTATTCGCCATTAACAGTCT 300  
QY 301 TTGAACCAATCTCTATCTCTGACCAAGACCATTTAAAGAAATGATTCATCGGTTT 360  
Db 301 TTGAACCAATCTCTATCTCTGACCAAGACCATTTAAAGAAATGATTCATCGGTTT 360  
QY 361 TGGCATACTACAGACAAAGAAATCTATGACAGCAAGGTATCAGAAACATATGCTC 420  
Db 361 TGGCATACTACAGACAAAGAAATCTATGACAGCAAGGTATCAGAAACATATGCTC 420  
QY 421 CATCATTCATCTGGGGGAATGTGGCCACTGCAGGAGATGCCACCTGAGTTAA 474  
Db 421 CATCATTCATCTGGGGGAATGTGGCCACTGCAGGAGATGCCACCTGAGTTAA 474

RESULT 2  
AX128507  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1  
Vogeli,G., Wood,L.S. and Merchant,K.  
G protein-coupled receptors expressed in brain  
Patent: WO 0131014-A 15 03-MAY-2001;  
PHARMACIA & UPJOHN COMPANY (US)  
FEATURES  
source  
Location/Qualifiers  
1..1191  
/organism="Homo sapiens"

CDS  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
1..1191  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAC38938.1"  
/db\_xref="GI:14134975"

/translation="MFRPLVNLSHIYFKFQYGVAPVRSCRPNTDGISSLENLLAS  
IIORFVWVSAVTCFNGNIVICMRPVIRSENKLYAMSIISLACADLMGILYFVIG  
FDLKEGYKNHQAOLMESTHCOLVSLAISTEVSLLITFLLEKYICIVVPRCV  
RPGKRTIVLIIWIGTGPVIAFIPLSNKEFKNYGTNGVCPPLSESTESIGAQIY  
SVAIFGILNLAFLIIVFSYGMFYSVHOSATIAPEIRNQVKEMILAKRFIFIFDALC  
WFPFVWVYIGYSAINSLNPIYITTPFKEMIHREWHNYRQKSMDSKIRKMLHSHSGKCGHCRCHLS"  
BASE COUNT 340 a 229 c 226 g 396 t  
ORIGIN

Query Match 92.7%; Score 439.2; DB 6; Length 1191;  
Best Local Similarity 96.8%; Pred. No. 1.1e-91;  
Matches 459; Conservative 0; Mismatches 13; Indels 2; Gaps 1;  
QY 1 GCCAGATTATTCAGTGGCAATTTCTTGGTATTAAATTTGGCGCATTTATCATATA 60  
Db 643 GCCAGATTATTCAGTGGCAATTTCTTGGTATTAAATTTGGCGCATTTATCATATA 702  
QY 61 GTTTTCTCTATGGAAGCATGTTTATATAGTGTTCATCAAAAGTCCATACAGCAACTGAA 120  
Db 703 GTTTTCTCTATGGAAGCATGTTTATATAGTGTTCATCAAAAGTCCATACAGCAACTGAA 762  
QY 121 ATACCGAATCAAGTAAAAAGAGATGATCTTCCCAAGCTTTTCTTATATAGTATT 180  
Db 763 ATACCGAATCAAGTAAAAAGAGATGATCTTCCCAAGCTTTTCTTATATAGTATT 822  
QY 181 ACTGATGATTATCTGGTATACCATTTTGTAGCGAAACCTCTTTCACTGCTTCAGGTA 240  
Db 823 ACTGATGATTATCTGGTATACCATTTTGTAGCGAAATTTCTTCACTGCTTCAGGTA 882  
QY 241 GAAATACAGGTACCAATACCTCTTGGGTAGTATGTTGTTATTCGCCATTAACAGTCT 300  
Db 883 GAAATACAGGTACCAATACCTCTTGGGTAGTATGTTGTTATTCGCCATTAACAGTCT 942  
QY 301 TTGAACCAATCTCTATCTCTGACCAAGACCATTTAAAGAAATGATTCATCGGTTT 360  
Db 943 TTGAACCAATCTCTATCTCTGACCAAGACCATTTAAAGAAATGATTCATCGGTTT 1002  
QY 361 TGGCATACTACAGACAAAGAAATCTATGACAGCAAGGTATCAGAAACATATGCTC 420  
Db 1003 TGGTATACTACAGACAAAGAAATCTATGACAGCAAGGTATCAGAAACATATGCTC 1060  
QY 421 CATCATTCATCTGGGGGAATGTGGCCACTGCAGGAGATGCCACCTGAGTTAA 474  
Db 1061 CATCATTCATCTGGGGGAATGTGGCCACTGCAGGAGATGCCACCTGAGTTAA 1114

RESULT 3  
AX549337  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1  
Burmer,G.C., Roush,C.L. and Brown,J.P.  
Antigenic peptides, such as for G protein-coupled receptors  
(GPCRs), antibodies thereto, and systems for identifying such  
antigenic peptides  
Patent: WO 02061087-A 622 08-AUG-2002;  
Lifespan Biosciences, Inc. (US)

## FEATURES

## Location/Qualifiers

## source

1..2274  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"

BASE COUNT 669 a 460 c 421 g 724 t  
ORIGIN

Query Match 92.7%; Score 439.2; DB 6; Length 2274;  
Best Local Similarity 96.8%; Pred. No. 1e-91;  
Matches 459; Conservative 0; Mismatches 13; Indels 2; Gaps 1;

QY 1 GCCCAGATTATTCAGTGGCAATTTCTGGTATTAATATGGCGCATTTATCATCATTA 60  
Db 1726 GCCCAGATTATTCAGTGGCAATTTCTGGTATTAATATGGCGCATTTATCATCATTA 1785  
QY 61 GTTTTTCCTATGGAAGCATGTTTATAGTGTTCATCAAAAGTGCCATAACAGCAACTGAA 120  
Db 1786 GTTTTTCCTATGGAAGCATGTTTATAGTGTTCATCAAAAGTGCCATAACAGCAACTGAA 1845  
QY 121 ATACGGAATCAAGTTAAAAAGAGATGATCCTTGGCCAAAGCTTTTCTTTATAGTATTT 180  
Db 1846 ATACGGAATCAAGTTAAAAAGAGATGATCCTTGGCCAAAGCTTTTCTTTATAGTATTT 1905  
QY 181 ACTGATGCATTAATGCTGATACCCATTTTGTAGAGCAACCTCTTTCACGCTTCAGGTA 240  
Db 1906 ACTGATGCATTAATGCTGATACCCATTTTGTAGTGAATTTCTTCACTGCTTCAGGTA 1965  
QY 241 GAATACAGAGTACCATACCTCTGGTAGTGATGGTATTCGCGCATTAACAGTGT 300  
Db 1966 GAATACAGAGTACCATACCTCTGGTAGTGATGGTATTCGCGCATTAACAGTGT 2025  
QY 301 TTGAACCCAAATCTCTATCTGACCAACAGACCATTTAAAGAAATGATTCATCGGTT 360  
Db 2026 TTGAACCCAAATCTCTATCTGACCAACAGACCATTTAAAGAAATGATTCATCGGTT 2085  
QY 361 TGGCATACTACAGACAAGAAATCTATGACAGCAAGAGTATCAGAAACATATGCTC 420  
Db 2086 TGGTATACTACAGACAAGAAATCTATGACAGCAAGAGTATCAGAAACATATGCTC 2143  
QY 421 CATCATTCATCTGGGGGAATGTGGCACTGCAGGAGATGCCACCTGAGTTAA 474  
Db 2144 CATCATTCATCTGGGGGAATGTGGCACTGCAGGAGATGCCACCTGAGTTAA 2197

RESULT 4  
AF190500 2274 bp mRNA linear PRI 10-AUG-2001  
LOCUS Homo sapiens leucine-rich repeat-containing G protein-coupled  
DEFINITION receptor 7 (LGR7) mRNA, complete cds.

ACCESSION AF190500  
VERSION AF190500.1 GI:10441729

KEYWORDS

SOURCE Homo sapiens (human)

## ORGANISM

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

1 (bases 1 to 2274)  
Hsu, S.-Y., Kudo, M., Chen, T., Nakabayashi, K., Bhalla, A., van der Spek, P.J., van Duin, M. and Hsueh, A.J.  
The three subfamilies of leucine-rich repeat-containing G protein-coupled receptors (LGR): identification of LGR6 and LGR7 and the signaling mechanism for LGR7  
Mol. Endocrinol. 14 (8), 1257-1271 (2000)

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

2 (bases 1 to 2274)  
Hsu, S.-Y. and Hsueh, A.J.W.

## AUTHORS

## TITLE

Direct Submission  
Submitted (28-SEP-1999) GYN/OB, Stanford University, MSOB S385,  
Stanford, CA 94305, USA

## JOURNAL

## LOCATION/Qualifiers

## source

1..2274

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

1..2274

/gene="LGR7"

1..2274

/gene="LGR7"

/codon\_start=1

/product="leucine-rich repeat-containing G protein-coupled

receptor 7"

/protein\_id="AAG17167.1"

/db\_xref="GI:10441730"

/translation="MTSGSVFFVILIFGKIFSHGGODVKCSLGYPPCGNITKCLPOL

LHCNVDDCGNADDEQNDGNGWSQPKYFASYKMTSQYFEATPECLVGSVPY

QICOGLEDCEDELENAVPSVSSNTAMSLQWNLIRKLPDPCFKYHDLQKLIQNN

KITSISIAFGLNSLKULSHNRITFLKGFVEDLHRLEWLIIEDNLSRISPTFL

YGLNSLILLLMMNVJTRPDLPCQHPRLHLDLGNHNLRLATFISCSNPTVL

VMKKNINLHNTFAPLQKDELQNGKIENPLFKDLKLSQLMSYNPQKI

QANQFDLVKLSLEGIEISNIQORMPRLMNLSHYFKFYQCYGAPHYVRSCKPN

TDQISLLENLASIQRVFWVVSATVCFGNIFVICMRPVIIRSENKLYAMSIIISCCA

DOIMGVILFVIGGFDLKFECEYNKHAQLPMSTHCOLVGSIALILSEVSVILLFELT

EXKICVYPRCPYRCKRIIVLILWITGFIIVAFIPLSNKEFFKNYGTNGVCFPL

HSEDTESIGAIYSVAIFGINIAFIIIVSYGSMFYSHVQSAITATIRNQVKEM

ILAKFFPIVFTDALCWPIFLVYKLSLLOVEIFCTISWVFIILPINSALNPLIYT

LLTRPFKEMIHFWNYRQRKSMDSKGOKTYAPSFIVWMPLOEMPPELMKPLFTY

PCMSLISOSTRLNSYS"

BASE COUNT 669 a 460 c 421 g 724 t

## ORIGIN

Query Match 92.7%; Score 439.2; DB 9; Length 2274;

Best Local Similarity 96.8%; Pred. No. 1e-91;

Matches 459; Conservative 0; Mismatches 13; Indels 2; Gaps 1;

QY 1 GCCCAGATTATTCAGTGGCAATTTCTGGTATTAATATGGCGCATTTATCATCATTA 60

Db 1726 GCCCAGATTATTCAGTGGCAATTTCTGGTATTAATATGGCGCATTTATCATCATTA 1785

QY 61 GTTTTTCCTATGGAAGCATGTTTATAGTGTTCATCAAAAGTGCCATAACAGCAACTGAA 120

Db 1786 GTTTTTCCTATGGAAGCATGTTTATAGTGTTCATCAAAAGTGCCATAACAGCAACTGAA 1845

QY 121 ATACGGAATCAAGTTAAAAAGAGATGATCCTTGGCCAAAGCTTTTCTTTATAGTATTT 180

Db 1846 ATACGGAATCAAGTTAAAAAGAGATGATCCTTGGCCAAAGCTTTTCTTTATAGTATTT 1905

QY 181 ACTGATGCATTAATGCTGATACCCATTTTGTAGAGCAACCTCTTTCACGCTTCAGGTA 240

Db 1906 ACTGATGCATTAATGCTGATACCCATTTTGTAGTGAATTTCTTCACTGCTTCAGGTA 1965

QY 241 GAATACAGAGTACCATACCTCTGGTAGTGATGGTATTCGCGCATTAACAGTGT 300

Db 1966 GAATACAGAGTACCATACCTCTGGTAGTGATGGTATTCGCGCATTAACAGTGT 2025

QY 301 TTGAACCCAAATCTCTATCTGACCAACAGACCATTTAAAGAAATGATTCATCGGTT 360

Db 2026 TTGAACCCAAATCTCTATCTGACCAACAGACCATTTAAAGAAATGATTCATCGGTT 2085

QY 361 TGGCATACTACAGACAAGAAATCTATGACAGCAAGAGTATCAGAAACATATGCTC 420

Db 2086 TGGTATACTACAGACAAGAAATCTATGACAGCAAGAGTATCAGAAACATATGCTC 2143

QY 421 CATCATTCATCTGGGGGAATGTGGCCACTGCAGGAGATGCCACCTGAGTTAA 474

Db 2144 CATCATTCATCTGGGGGAATGTGGCCACTGCAGGAGATGCCACCTGAGTTAA 2197

## RESULT 5

## LOCUS

## BD135245

## DEFINITION

## Novel mammalian G protein-coupled

## leucine-rich repeating domain.

## ACCESSION

## BD135245

## VERSION

## BD135245.1

## GI:23230190

2467 bp

DNA

linear

PAT 18-SEP-2002

KEYWORDS JP 2002507406-A/3.  
 SOURCE Homo sapiens (human)  
 ORGANISM  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE  
 AUTHORS Hsueh,A.J.W., Hsu,S.Y., Liang,S.G. and Spe,P.J.V.D.  
 TITLE Novel mammalian G protein-coupled receptor having extracellular  
 leucine-rich repeating domain  
 JOURNAL Patent: JP 2002507406-A 3 12-MAR-2002;  
 THE BOARD OF TRUSTEES OF THE LELAND STANFORD JUNIOR UNIVERSITY,  
 AKZO NOBEL NV  
 COMMENT OS Homo sapiens (human)  
 PN JP 2002507406-A/3  
 PD 12-MAR-2002  
 PR 25-MAR-1999 JP 2000537903  
 PR 26-MAR-1998 US 60/079501  
 PI AARON J W HSUEH, SHEAU YU HSU, SHAN GUANG LIANG, PETRUS JOHANNES  
 PI VAN DER SPEK  
 PC C12N15/09,A01K67/027,C07K14/705,C07K16/28,C12N1/15,C12N1/19,  
 PC C12N5/10,C12P21/08,C12N15/00,C12N5/00  
 CC Novel mammalian G protein-coupled receptor having CC  
 extracellular  
 CC leucine-rich repeating domain  
 FH Key Location/Qualifiers  
 FT source 1..2467  
 FT /organism="Homo sapiens (human)"  
 FT /mol\_type="genomic DNA"  
 FT /db\_xref="taxon:9606"  
 FEATURES  
 source 1..2467  
 Location/Qualifiers  
 1..2467  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 BASE COUNT 747 a 487 c 474 g 759 t  
 ORIGIN  
 Query Match 92.7%; Score 439.2; DB 6; Length 2467;  
 Best Local Similarity 96.8%; Pred. No. 1e-91;  
 Matches 459; Conservative 0; Mismatches 13; Indels 2; Gaps 1;  
 QY 1 GCCAGATTATTTCAGTGGCAATTTCTTGTTATTAATTTGGCCGCAATTTATCATATA 60  
 DB 1919 GCCAGATTATTTCAGTGGCAATTTCTTGTTATTAATTTGGCCGCAATTTATCATATA 1978  
 QY 61 GTTTTTCCTATGGAAGCAATTTTATAGTGTTCATCAAAAGTGCCTTAACAGCACTGAA 120  
 DB 1979 GTTTTTCCTATGGAAGCAATTTTATAGTGTTCATCAAAAGTGCCTTAACAGCACTGAA 2038  
 QY 121 ATACGGAATCAAGTAAAGAGATGATCCTTGCCAAAGCTTTTCTTTTATAGTATT 180  
 DB 2039 ATACGGAATCAAGTAAAGAGATGATCCTTGCCAAAGCTTTTCTTTTATAGTATT 2098  
 QY 181 ACTGATGCAATATGCTGGAPACCAATTTTGTAGCGAAACCTCTTCACTGCTTCAGTA 240  
 DB 2099 ACTGATGCAATATGCTGGAPACCAATTTTGTAGCGAAACCTCTTCACTGCTTCAGTA 2158  
 QY 241 GAAATACAGGTACCAATCCTTGGGAGTGGTATTCTGCAATTAACAGTCT 300  
 DB 2159 GAAATACAGGTACCAATCCTTGGGAGTGGTATTCTGCAATTAACAGTCT 2218  
 QY 301 TTGAACCAATTCCTATCTACTCTGACCAAGACCAATTTAAAGAAATGATTCATCGGTT 360  
 DB 2219 TTGAACCAATTCCTATCTACTCTGACCAAGACCAATTTAAAGAAATGATTCATCGGTT 2278  
 QY 361 TGCATTAACAGCAAGAAATCTATGAGCAAGCAAGTATGCAAAACATATGCTC 420  
 DB 2279 TGCATTAACAGCAAGAAATCTATGAGCAAGCAAGTATGCAAAACATATGCTC 2336  
 QY 421 CATCATTCATCTGGGGGAATCTGCCACAGCAGATGCCACCTGAGTTAA 474  
 DB 2337 CATCATTCATCTGGGGGAATCTGCCACAGCAGATGCCACCTGAGTTAA 2390

RESULT 6  
 BD135246  
 LOCUS  
 DEFINITION Novel mammalian G protein-coupled receptor having extracellular  
 leucine-rich repeating domain.  
 ACCESSION BD135246  
 VERSION BD135246.1 GI:23230191  
 KEYWORDS JP 2002507406-A/4.  
 SOURCE Homo sapiens (human)  
 ORGANISM  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE  
 AUTHORS Hsueh,A.J.W., Hsu,S.Y., Liang,S.G. and Spe,P.J.V.D.  
 TITLE Novel mammalian G protein-coupled receptor having extracellular  
 leucine-rich repeating domain  
 JOURNAL Patent: JP 2002507406-A 4 12-MAR-2002;  
 THE BOARD OF TRUSTEES OF THE LELAND STANFORD JUNIOR UNIVERSITY,  
 AKZO NOBEL NV  
 COMMENT OS Homo sapiens (human)  
 PN JP 2002507406-A/4  
 PD 12-MAR-2002  
 PR 25-MAR-1999 JP 2000537903  
 PR 26-MAR-1998 US 60/079501  
 PI AARON J W HSUEH, SHEAU YU HSU, SHAN GUANG LIANG, PETRUS JOHANNES  
 PI VAN DER SPEK  
 PC C12N15/09,A01K67/027,C07K14/705,C07K16/28,C12N1/15,C12N1/19,  
 PC C12N5/10,C12P21/08,C12N15/00,C12N5/00  
 CC Novel mammalian G protein-coupled receptor having CC  
 extracellular  
 CC leucine-rich repeating domain  
 FH Key Location/Qualifiers  
 FT source 1..3584  
 FT /organism="Homo sapiens (human)"  
 FT /mol\_type="genomic DNA"  
 FT /db\_xref="taxon:9606"  
 FEATURES  
 source 1..3584  
 Location/Qualifiers  
 1..3584  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 BASE COUNT 1124 a 670 c 647 g 1139 t 4 others  
 ORIGIN  
 Query Match 92.7%; Score 439.2; DB 6; Length 3584;  
 Best Local Similarity 96.8%; Pred. No. 9.8e-92;  
 Matches 459; Conservative 0; Mismatches 13; Indels 2; Gaps 1;  
 QY 1 GCCAGATTATTTCAGTGGCAATTTCTTGTTATTAATTTGGCCGCAATTTATCATATA 60  
 DB 1744 GCCAGATTATTTCAGTGGCAATTTCTTGTTATTAATTTGGCCGCAATTTATCATATA 1803  
 QY 61 GTTTTTCCTATGGAAGCAATTTTATAGTGTTCATCAAAAGTGCCTTAACAGCACTGAA 120  
 DB 1804 GTTTTTCCTATGGAAGCAATTTTATAGTGTTCATCAAAAGTGCCTTAACAGCACTGAA 1863  
 QY 121 ATACGGAATCAAGTAAAGAGATGATCCTTGCCAAAGCTTTTCTTTTATAGTATT 180  
 DB 1864 ATACGGAATCAAGTAAAGAGATGATCCTTGCCAAAGCTTTTCTTTTATAGTATT 1923  
 QY 181 ACTGATGCAATATGCTGGATACCAATTTTGTAGCGAAACCTCTTCACTGCTTCAGTA 240  
 DB 1924 ACTGATGCAATATGCTGGATACCAATTTTGTAGCGAAACCTCTTCACTGCTTCAGTA 1983  
 QY 241 GAAATACAGGTACCAATCCTTGGGAGTGGTATTCTGCAATTAACAGTCT 300  
 DB 1984 GAAATACAGGTACCAATCCTTGGGAGTGGTATTCTGCAATTAACAGTCT 2043  
 QY 301 TTGAACCAATTCCTATCTACTCTGACCAAGACCAATTTAAAGAAATGATTCATCGGTT 360  
 DB 2044 TTGAACCAATTCCTATCTACTCTGACCAAGACCAATTTAAAGAAATGATTCATCGGTT 2103  
 QY 361 TGCATTAACAGCAAGAAATCTATGAGCAAGCAAGTATGCAAAACATATGCTC 420



Db 2104 TGGTAACTACAGACAAAGAAATCTATGGACACAAAGG--TCAGAAACATATGCTC 2161

QY 421 CATCATTCATCTGGGGGAAATGTGCCACTGCAGGAGATGCCACCTGAGTTAA 474  
|||||

Db 2162 CATCATTCATCTGGGTGAAATGTGCCACTGCAGGAGATGCCACCTGAGTTAA 2215  
|||||

RESULT 7  
AC107219  
LOCUS 187431 bp DNA linear PRI 27-MAR-2002  
DEFINITION Homo sapiens BAC clone RP11-575B4 from 4, complete sequence.  
AC107219  
VERSION AC107219.5 GI:19526153  
KEYWORDS HTG.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE  
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
TITLE Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
JOURNAL 1 (bases 1 to 187431)  
MEDLINE Sulston, J.E. and Waterston, R.  
Toward a complete human genome sequence  
99063792 Genome Res. 8 (11), 1097-1108 (1998).  
9847074  
2 (bases 1 to 187431)  
Isak, A., Meyer, R. and Boyer, E.  
The sequence of Homo sapiens BAC clone RP11-575B4  
Unpublished (2001)  
3 (bases 1 to 187431)  
Waterston, R.H.  
Direct Submission  
Submitted (16-JAN-2002) Genome Sequencing Center, Washington  
University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
MO 63108, USA  
4 (bases 1 to 187431)  
Waterston, R.H.  
Direct Submission  
Submitted (21-FEB-2002) Genome Sequencing Center, Washington  
University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
MO 63108, USA  
5 (bases 1 to 187431)  
Waterston, R.H.  
Direct Submission  
Submitted (16-MAR-2002) Genome Sequencing Center, Washington  
University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
MO 63108, USA  
6 (bases 1 to 187431)  
Waterston, R.H.  
Direct Submission  
Submitted (27-MAR-2002) Department of Genetics, Washington  
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
On Mar 16, 2002 this sequence version replaced gi:18855160.  
----- Genome Center  
Center: Washington University Genome Sequencing Center  
Center code: WUGSC  
Web site: <http://genome.wustl.edu/gsc>  
Contact: [sapiens@watson.wustl.edu](mailto:sapiens@watson.wustl.edu)  
----- Summary Statistics  
Center project name: H\_NH0575B04  
-----

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:  
Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:  
The RP11-11 human BAC library was made from the blood of one male donor, as described by Osogawa, K., Woon, P. Y., Zhao, B., Frengen, E., Tateno, M., Catanese, J. J. and de Jong, P. J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>  
VECTOR: pBACE3.6

NEIGHBORING SEQUENCE INFORMATION:  
The clone sequenced to the left is CTD-2005A22; the clone sequenced to the right is RP11-563M18. Actual start of this clone is at base position 1 of RP11-575B4; actual end is at base position 187431 of RP11-575B4.

Sequence derived from one plasmid subclone, base position 102011 to 102054.

FEATURES	Location/Qualifiers
source	1..187431
	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
	/chromosome="4"
	/map="4"
	/clone="RP11-575B4"
	/clone_lib="RPC1-11"
repeat_region	1..408
	/rpt_family="L1"
repeat_region	563..697
	/rpt_family="Alu"
repeat_region	829..938
	/rpt_family="Alu"
repeat_region	1227..1326
	/rpt_family="MIR"
repeat_region	1327..2602
	/rpt_family="L1"
repeat_region	2973..3009
	/rpt_family="AT-rich"
repeat_region	3233..3534
	/rpt_family="Alu"
repeat_region	3582..3616
	/rpt_family="AT-rich"
repeat_region	3685..3772
	/rpt_family="L2"
repeat_region	4141..4297
	/rpt_family="MaLR"
repeat_region	4298..4593
	/rpt_family="Alu"
repeat_region	4594..4804
	/rpt_family="MaLR"
repeat_region	5020..5436
	/rpt_family="MaLR"
repeat_region	5483..5506
	/rpt_family="AT-rich"
repeat_region	6006..6170
	/rpt_family="MaLR"
repeat_region	6709..6748
	/rpt_family="L2"
repeat_region	6863..7446
	/rpt_family="L1"
repeat_region	7454..8078
	/rpt_family="ERV1"
repeat_region	8079..8175
	/rpt_family="L1"
repeat_region	8176..8477

Db	13263	GTCTTCATCAAAAGTCCCATACAGCAACTGAAATACAGCAATCAAGTTAAAAAGAGATGA	13382
QY	149	TCCTTCCCAAAGCTTTTCTTTATAGTATTTACTGATGCTATTATGCTGATACCATTT	208
Db	13323	TCCTTCCCAAAGCTTTTCTTTATAGTATTTACTGATGCTATTATGCTGATACCATTT	13382
QY	209	TTGTACGGAACCTCTTTCCTACTGCTTCAGGTAGAGAAATACAGGTACCATTA	258
Db	13383	TTGTAGTGAATTTCTTTCACTGCTTCAGGTAGAGAAATACAGGTACCATTA	13432
RESULT 8			
LOCUS	B0141178		
DEFINITION	Novel G protein-coupled receptor protein and its DNA.		
ACCESSION	B0141178		
VERSION	B0141178.1	GI:323236123	
KEYWORDS	WO 0204640-A/7		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	1 (bases 1 to 1065)		
AUTHORS	Moriya,T., Ito,T., Shintani,Y. and Miyajima,N.		
TITLE	Novel G protein-coupled receptor protein and its DNA		
JOURNAL	Patent: WO 0204640-A 7 17-JAN-2002;		
COMMENT	TAKEEDA CHEMICAL INDUSTRIES LTD, TAKEO MORIYA, TAKASHI ITO, YASUSHI SHINTANI, NOBUYUKI MIYAJIMA		
	OS Homo sapiens (human)		
	PN WO 0204640-A/7		
	PD 17-JAN-2002		
	PF 06-JUL-2001 WO 2001JP005878		
	PR 07-JUL-2000 JP 00P 211389, 18-DEC-2000 JP 00P 383794		
	TAKEO MORIYA, TAKASHI ITO, YASUSHI SHINTANI, NOBUYUKI MIYAJIMA		
	CI2N15/12, CI2N17/21, CO7K14/705, CO7K16/28, CI2P21/02, CI2Q1/68, PC		
	A61K38/00,		
	PC A61K45/00, A61K48/00, A61P1/00, A61P3/00, A61P9/00, A61P25/28, PC		
	A61P29/00,		
	PC A61P35/00, A61P37/00, G01N33/15, G01N33/50, G01N33/53, G01N33/566//		
	PC (CI2N1/21, CI2R1:19), (CI2P21/02, CI2R1:19)		
	CC Novel G protein-coupled receptor protein and its DNA. FH Key		
	FT source	1..1065	
	FT Location/Qualifiers	/organism='Homo sapiens (human)'	
FEATURES			
source	1..1065		
	/organism="Homo sapiens"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:9606"		
BASE COUNT	277 a 214 c 213 g 361 t		
ORIGIN			
Query Match	36.2%;	Score 171.8;	DB 6; Length 1065;
Best Local Similarity	67.4%;	Pred. No. 1.2e-29;	
Matches	242;	Conservative 0;	Mismatches 117; Indels 0; Gaps 0;
QY	10	TATTCAGTGGCAATTTTCTTGTGTATTAATTTGGCCCACTTATCATCATATGTTTTC	69
Db	568	TATTCCTTGGAAATTTCTAGTGTGAACCTGTGCTGCTTTCTCATCATTTGTTTTC	627
QY	70	TATGGAAGCATGTTTATAGTGTTCATCAACTGCCATACACCACTGGAATACGAAT	129
Db	628	TATATTATGTTCTGTTCCATTTCAAHAACCCCTTCGACCCACAGAAATAGGAAT	687
QY	130	CAAGTTAAAAAGAGATGATCCTTCGCAAGCTTTTCTTTATAGTATTTACTGATGCA	189
Db	688	TGTTTGGAGAGAGGTGGTGTGCAATCGTTCTTTTATAGTCTCTCTGATGCC	747
QY	190	TTATGCTGGATACCATTTTGTAGGAAACCTCTTTCACCTGCTTACGTTAGAAATACA	249
Db	748	ATCTGCTGGATTCCTGTATTATTAGTTAGTTAAATCTTTCCCTCTTCCGGGTGGAATACA	807

QY 250 GGTACCATACCTCTGGTAGTGGTATTTCTCCATTAAACAGCTGTTGAACCCA 309  
 Db 808 GACACATGACTCTCTGGATAGTATTTTCTTCCAGTTAAACAGTCTTGAATCCA 867  
 QY 310 ATTCTTATCTCTGACCAACAGACCATTTAAAGAAATGATTCATCGGTTTTGGCATPA 368  
 Db 868 ATCCTCTATCTCTCACACCAACCTTTTAAAGACAAAGTTGAACAGCTGCTGCACAA 926

RESULT 9  
 AXI48176  
 LOCUS 1068 bp DNA linear PAT 08-JUN-2001  
 DEFINITION Sequence 17 from Patent WO0136471.  
 ACCESSION AXI48176  
 VERSION AXI48176.1 GI:14347081  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1  
 AUTHORS Chen, R., Dang, H. T. and Lowitz, K. P.  
 TITLE Endogenous and non-endogenous versions of human g protein-coupled receptors  
 JOURNAL Patent: WO 0136471-A 17 25-MAY-2001;  
 Arena Pharmaceuticals, Inc. (US)  
 FEATURES  
 source Location/Qualifiers  
 1..1068  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"

BASE COUNT 278 a 215 c 214 g 361 t  
 ORIGIN

Query Match 36.2%; Score 171.8; DB 6; Length 1068;  
 Best Local Similarity 67.4%; Pred. No. 1.2e-29;  
 Matches 242; Conservative 0; Mismatches 117; Indels 0; Gaps 0;

QY 10 TATTCAGTGGCAATTTTCTTGTTAAATTTGGCCGCAATTTATCATCATAGTTTTC 69  
 Db 568 TATTCCTCTGGATTTTCTAGTGTGACTGTGCTGCTTCTCATCATTTGTGTTC 627  
 QY 70 TATGGAAGCATGTTTATAGTGTTCATCAAGTGCCTACACAGCAACTGAAATACGGAAT 129  
 Db 628 TATATTACTATGTCTGTTCCATTCAAAAACCCCTTGACAGACCACAGAAAGTAAGAAAT 687  
 QY 130 CAAGTATAAAGACAGATGATCTCTGCAAGTTCATCAAGTGCCTACACAGCAACTGAAATACGGAAT 129  
 Db 688 TGTTTGGAGAGAGAGTGGCTGTTCGAATCGTTCTTTTATAGTGTCTCTGATGCC 747  
 QY 190 TTATGCTGGATACCCATTTTGTAGCGAAACCTCTTTCACGTTCAGGTAGAAATACCA 249  
 Db 748 ATCTGCTGGATTCCTGATTTAGTTAAATCCTTTCCCTCTCCGGGTGGAATACCA 807  
 QY 250 GGTACCATACCTCTGGTAGTGGTATTTCTCCATTAAACAGCTGTTGAACCCA 309  
 Db 808 GACACATGACTCTCTGGATAGTATTTTCTTCCAGTTAAACAGTCTTGAATCCA 867  
 QY 310 ATTCTTATCTCTGACCAACAGACCATTTAAAGAAATGATTCATCGGTTTTGGCATPA 368  
 Db 868 ATCCTCTATCTCTCACACCAACCTTTTAAAGACAAAGTTGAACAGCTGCTGCACAA 926

RESULT 10  
 BD141176  
 LOCUS 1473 bp DNA linear PAT 18-SEP-2002  
 DEFINITION Novel G protein-coupled receptor protein and its DNA.  
 ACCESSION BD141176  
 VERSION BD141176.1 GI:23236121  
 KEYWORDS WO 0204640-A/5.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 1473)  
 AUTHORS Moriya, T., Ito, T., Shintani, Y. and Miyajima, N.  
 TITLE Novel G protein-coupled receptor protein and its DNA  
 JOURNAL Patent: WO 0204640-A 5 17-JAN-2002;  
 TAKEDA CHEMICAL INDUSTRIES LTD, TAKEO MORIYA, TAKASHI ITO, YASUSHI SHINTANI, NOBUYUKI MIYAJIMA  
 COMMENT OS Homo sapiens (human)  
 PN WO 0204640-A/5  
 PD 17-JAN-2002  
 PF 06-JUL-2001 WO 2001JP005878  
 PR 07-JUL-2000 JP OOP 211989, 18-DEC-2000 JP OOP 383794 PI  
 TAKEO MORIYA, TAKASHI ITO, YASUSHI SHINTANI, NOBUYUKI MIYAJIMA PC  
 C12N15/12, C12N1/21, C07K14/705, C07K16/28, C12P21/02, C12Q1/68, PC  
 A61K38/00,  
 PC A61K45/00, A61K48/00, A61P1/00, A61P3/00, A61P9/00, A61P25/28, PC  
 A61P29/00,  
 PC A61P35/00, A61P37/00, G01N33/15, G01N33/50, G01N33/53, G01N33/566//  
 PC (C12N1/21, C12R1:19), (C12P21/02, C12R1:19)  
 CC Novel G protein-coupled receptor protein and its DNA. FH Key  
 Location/Qualifiers  
 FT source 1..1473  
 FT /organism="Homo sapiens (human)".  
 FEATURES  
 source Location/Qualifiers  
 1..1473  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"

BASE COUNT 405 a 310 c 278 g 480 t  
 ORIGIN

Query Match 36.2%; Score 171.8; DB 6; Length 1473;  
 Best Local Similarity 67.4%; Pred. No. 1.2e-29;  
 Matches 242; Conservative 0; Mismatches 117; Indels 0; Gaps 0;

QY 10 TATTCAGTGGCAATTTTCTTGTTAAATTTGGCCGCAATTTATCATCATAGTTTTC 69  
 Db 976 TATTCCTCTGGATTTTCTAGTGTGACTGTGCTGCTTCTCATCATTTGTGTTC 1035  
 QY 70 TATGGAAGCATGTTTATAGTGTTCATCAAGTGCCTACACAGCAACTGAAATACGGAAT 129  
 Db 1036 TATATTACTATGTCTGTTCCATTCAAAAACCCCTTGACAGACCACAGAAAGTAAGAAAT 1095  
 QY 130 CAAGTATAAAGACAGATGATCTCTGCAAGTTCATCAAGTGCCTACACAGCAACTGAAATACGGAAT 129  
 Db 1096 TGTTTGGAGAGAGAGTGGCTGTTCGAATCGTTCTTTTATAGTGTCTCTGATGCC 1155  
 QY 190 TTATGCTGGATACCCATTTTGTAGCGAAACCTCTTTCACGTTCAGGTAGAAATACCA 249  
 Db 1156 ATCTGCTGGATTCCTGATTTAGTTAAATCCTTTCCCTCTCCGGGTGGAATACCA 1215  
 QY 250 GGTACCATACCTCTGGTAGTGGTATTTCTCCATTAAACAGCTGTTGAACCCA 309  
 Db 1216 GACACATGACTCTCTGGATAGTATTTTCTTCCAGTTAAACAGTCTTGAATCCA 1275  
 QY 310 ATTCTTATCTCTGACCAACAGACCATTTAAAGAAATGATTCATCGGTTTTGGCATPA 368  
 Db 1276 ATCCTCTATCTCTCACACCAACCTTTTAAAGACAAAGTTGAACAGCTGCTGCACAA 1334

RESULT 11  
 BD141175  
 LOCUS 1545 bp DNA linear PAT 18-SEP-2002  
 DEFINITION Novel G protein-coupled receptor protein and its DNA.  
 ACCESSION BD141175  
 VERSION BD141175.1 GI:23236120  
 KEYWORDS WO 0204640-A/4.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 1545)  
 AUTHORS Moriya, T., Ito, T., Shintani, Y. and Miyajima, N.

TITLE	Novel G protein-coupled receptor protein and its DNA									
JOURNAL	Patent: WO 0204640-A 4 17-JAN-2002; TAKEDA CHEMICAL INDUSTRIES LTD, TAKEO MORIYA, TAKASHI ITO, YASUSHI SHINTANI, NOBUYUKI MIYAJIMA									
COMMENT	OS Homo sapiens (human) PN WO 0204640-A/4 PD 17-JAN-2002 PF 06-JUL-2001 WO 2001P005878 PI 07-JUL-2000 JP 00P 211989, 18-DEC-2000 JP 00P 383794 PI TAKEO MORIYA, TAKASHI ITO, YASUSHI SHINTANI, NOBUYUKI MIYAJIMA PC C12N15/12, C12N1/21, C07K14/705, C07K16/28, C12P21/02, C12P1/68, PC A61K38/00, PC A61K45/00, A61K48/00, A61P1/00, A61P3/00, A61P9/00, A61P25/28, PC A61P29/00, PC A61P35/00, A61P37/00, G01N33/15, G01N33/50, G01N33/56, // PC (C12N1/21, C12R1:19), (C12P21/02, C12R1:19) CC Novel G protein-coupled receptor protein and its DNA. FH Key Location/Qualifiers FT source 1. .1545 FT /organism='Homo sapiens (human)'. Location/Qualifiers 1. .1545 /organism='Homo sapiens' /mol_type='genomic DNA' /db_xref='taxon:9606'									
BASE COUNT	429 a 320 c 290 g 506 t									
ORIGIN										
Query Match	36.2%; Score 171.8; DB 6; Length 1545;									
Best Local Similarity	67.4%; Pred. No. 1.2e-29;									
Matches 242;	Conservative 0; Mismatches 117; Indels 0; Gaps 0;									
QY	10 TATTAGTGGCAATTTTCTTGGTATTAAATTTGGCGCATTTTGGCGCATTTATCATCATGTTTTTCC 69 									
Db	1048 TATTCCTGGAAATTTCTTAGGTGGAACCTTGTGCGCTTTTCTCATCATTTGTTTTCC 1107 									
QY	70 TAGGAGACATGTTTATAGTGTTCATCAAGTGCATTAACGACACTGAATACGGAAAT 129 									
Db	1108 TATATTACTATGTTCTGTTCATTCAAAAACGGCTTGCAGACACACAGAAATAGGAAT 1167 									
QY	130 CAAGTTAAAAAGAGATGATCCTTTGGCCAAACGTTTTTCTTTATAGTATTACTGTGCA 189 									
Db	1168 TGTTTTGGAGAGAGGTGGCTTTGTGCATCTGTTCTTTTATAGTGTCTCTGATGCC 1227 									
QY	190 TTATGCTGGATACCCATTTTGTAGGAAACCTCTTTTCAGTGTTCAGGTAGAAATACCA 249 									
Db	1228 ATCTGTGGATTCCTGTATTTGTAGTAAAACTCTTCCCTTCCGGGTGGAAATACCA 1287 									
QY	250 GGTACCATTAACCTCTTGGGTAGTATGGTATTCTGCGCATTAACAGTGTCTTGAACCCA 309 									
Db	1288 GACACATGCTCTCTGGATAGTGAATTTTTCCTTCAGTTACAGTGTCTTGAATCCA 1347 									
QY	310 ATTCTCTATACTCTGACCACAGACCAATTTAAGAAATGATTCATCGGTTTGGCAATAA 368 									
Db	1348 ATCTCTATACTCTCACAAACCACTTTTAAAGGCAAGTTGAAACAGCTGCTGCACAA 1406 									
RESULT 12										
BD141172	LOCUS BD141172 1830 bp DNA linear PAT 18-SEP-2002									
DEFINITION	Novel G protein-coupled receptor protein and its DNA.									
ACCESSION	BD141172									
VERSION	BD141172.1 GI:23236117									
KEYWORDS	WO 0204640-A/1.									
SOURCE	Homo sapiens (human)									
ORGANISM	Homo sapiens									
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 1830)									
AUTHORS	Moriya, T., Ito, T., Shintani, Y. and Miyajima, N.									
TITLE	Novel G protein-coupled receptor protein and its DNA									
JOURNAL	Patent: WO 0204640-A 17-JAN-2002; TAKEDA CHEMICAL INDUSTRIES LTD, TAKEO MORIYA, TAKASHI ITO, YASUSHI									

```
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
1..2049
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD28590.1"
/db_xref="GI:19578161"
/translation="MTVFLVFKHFLSLMTFLHFLHFLVFLVLYNKFALTGSGMIRPS
CQGFPCGNLTKCFRAPHGDDGKDDGNGADENGDTSGWATIFGVHGNANVAL
TQCFKQPCQDCRTECEVGNGLKSPMISNNVTLISLKKNKIHSPDKVFKY
TKLFLPQHCNIRHISKRFAFGCLNLIQILIDNPITRISORLFTGLNSFTLSMWN
NYLALPQKQAMQPOLNVLGNRIKYLNTGSLSCDVLDSNLTTELSPHL
FKDLKLDLIERIEIPNIRTPQMKNLSHYFNFRYCSYAPHVRIQMPITDGI
SSPDLANNILRIFWVLIARTFCFGLNFVGMERSPIKAENTHAMSFKILCCADCLM
GVILFFVPIIDIKYQVQKVALLMESVQCRMLGMLSTEVSLLVILYILEFL
VTFPSNPKRQTSVILICIMAGFLIAPFNWKDIFGNFGNGVGYFLIYDQ
TEDGSGKSLGIFGVLNLAFLIIVFSYITMFCISOKTALQOTVEVRCNFGREYAVN
RPFPIVSDAICWIPVYVKILSLFRVEIPDPTMSWIVFFLPVNSALNPILYLTIN
PFKDLKQLLKHQRKSTFKIKKSLSTSIVWIEDSSSLKGLVKNKITLGSIMKPV
"
sig_peptide 1..108
BASE COUNT 579 a 415 c 395 g 660 t
ORIGIN
Query Match 36.2%; Score 171.8; DB 6; Length 2049;
Best Local Similarity 67.4%; Pred. No. 1.1e-29;
Matches 242; Conservative 0; Mismatches 117; Indels 0; Gaps 0;
Qy 10 TATTCAGTGGCAATTTTCTGTGTATTAATTGGCCGCAATTATCATCATAGTTTTC 69
Dy 1549 TATTCCTGTGGAATTTTCTGTGTATTAATTGGCCGCAATTATCATCATAGTTTTC 1608
Qy 70 TATGGAACGATTTTATAGTGTTCATCAAGTGCATACAGCAACTGAATACGGAAT 129
Dy 1609 TATATATAGTGTTCCTGATCAATCAAAACCGCCTTCGAGACCAAGTAAGGAAT 1668
Qy 130 CAAGTTAAAAAGAGATGATCCTTGCCAAAGTTTCTTCTTATAGTATTTACTGATCA 189
Dy 1669 TGTTTTGGAGAGAGGTGGCTGTGCAAAATGCTTTCTTTTATAGTGTCTCTGATCC 1728
Qy 190 TTATGCTGGATACCAATTTTGTAGCGAAACCTCTTTCAGTGTTCAGTAGAATACCA 249
Dy 1729 ATCTGCTGGATTCCTGTATTTAGTAAATCTTTCCCTCTCCGGGTGGAATACCA 1788
Qy 250 GGTACCAATACCTCTCGGAGTAGTGTTCATCAAGTGCATACAGCAACTGAATACCA 309
Dy 1789 GACACAATGATCTCTGATAGTGTATTTCTCCATTAACAGCAACTGCTGACCA 1848
Qy 310 ATCTCTATCTCTGACCAACCAATTTAAAGAAATGATTCATCGTTTGGCATAA 368
Dy 1849 ATCCCTATATCTCTCACAACTTTTAAAGCAAGTTGAACAGCTGCTGCACAA 1907
RESULT 14
AX451566
LOCUS 2142 bp DNA linear PAT 03-JUL-2002
DEFINITION Sequence 5 from Patent WO0226824.
ACCESSION AX451566
VERSION AX451566.1 GI:21698550
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Baughn,M.R., Tribouley,C.M., Nguyen,D.B., Thornton,M., Yao,M.G.,
Kallik,D.A., Gandhi,A.R., Walla,N.K., Arvizu,C., Elliott,V.S.,
Hafalia,A.J., Ramkumar,J., Pei,J., Tang,Y.T., Yue,H., Reddy,R.,
Burford,N., Lu,D.A., Graul,R.C., Khan,F.A., Walsh,R.T., Ison,C.H.,
Richardson,T.W., Griffin,J.A., Warren,B.A., Yang,J., Lee,E.A. and
Harland,L.
G-protein coupled receptors
Patent: WO 02063004-A 49 15-AUG-2002;
Incyte Genomics, Inc. (US)
FEATURES
Location/Qualifiers
1..2181
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: 7485090CBI"
BASE COUNT 635 a 451 c 419 g 676 t
ORIGIN
Query Match 36.2%; Score 171.8; DB 6; Length 2181;
Best Local Similarity 67.4%; Pred. No. 1.1e-29;
Matches 242; Conservative 0; Mismatches 117; Indels 0; Gaps 0;
Qy 10 TATTCAGTGGCAATTTTCTGTGTATTAATTGGCCGCAATTATCATCATAGTTTTC 69
Dy 1681 TATTCCTGTGGAATTTTCTGTGTATTAATTGGCCGCAATTATCATCATAGTTTTC 1740
```

```
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
613 a 435 c 408 g 686 t
BASE COUNT
ORIGIN
Query Match 36.2%; Score 171.8; DB 6; Length 2142;
Best Local Similarity 67.4%; Pred. No. 1.1e-29;
Matches 242; Conservative 0; Mismatches 117; Indels 0; Gaps 0;
Qy 10 TATTCAGTGGCAATTTTCTGTGTATTAATTGGCCGCAATTATCATCATAGTTTTC 69
Dy 1642 TATTCCTGTGGAATTTTCTGTGTATTAATTGGCCGCAATTATCATCATAGTTTTC 1701
Qy 70 TATGGAACGATTTTATAGTGTTCATCAAGTGCATACAGCAACTGAATACGGAAT 129
Dy 1702 TATATATAGTGTTCCTGATCAATCAAAACCGCCTTCGAGACCAAGTAAGGAAT 1761
Qy 130 CAAGTTAAAAAGAGATGATCCTTGCCAAAGTTTCTTCTTATAGTATTTACTGATCA 189
Dy 1762 TGTTTTGGAGAGAGGTGGCTGTGCAATGCTTTCTTTTATAGTGTCTCTGATGCC 1821
Qy 190 TTATGCTGGATACCAATTTTGTAGCGAAACCTCTTTCAGTGTTCAGTAGAATACCA 249
Dy 1822 ATCTGCTGGATTCCTGTATTTAGTAAATCTTTCCCTCTCCGGGTGGAATACCA 1881
Qy 250 GGTACCAATACCTCTCGGAGTAGTGTTCATCAAGTGCATACAGCAACTGAATACCA 309
Dy 1892 GACACAATGATCTCTGATAGTGTATTTCTCCATTAACAGCAACTGCTGACCA 1941
Qy 310 ATCTCTATCTCTGACCAACCAATTTAAAGAAATGATTCATCGTTTGGCATAA 368
Dy 1942 ATCCCTATATCTCTCACAACTTTTAAAGCAAGTTGAACAGCTGCTGCACAA 2000
RESULT 15
AX709190
LOCUS 2181 bp DNA linear PAT 04-APR-2003
DEFINITION Sequence 49 from Patent WO02063004.
ACCESSION AX709190
VERSION AX709190.1 GI:29564784
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Baughn,M.R., Tribouley,C.M., Nguyen,D.B., Thornton,M., Yao,M.G.,
Kallik,D.A., Gandhi,A.R., Walla,N.K., Arvizu,C., Elliott,V.S.,
Hafalia,A.J., Ramkumar,J., Pei,J., Tang,Y.T., Yue,H., Reddy,R.,
Burford,N., Lu,D.A., Graul,R.C., Khan,F.A., Walsh,R.T., Ison,C.H.,
Richardson,T.W., Griffin,J.A., Warren,B.A., Yang,J., Lee,E.A. and
Harland,L.
G-protein coupled receptors
Patent: WO 02063004-A 49 15-AUG-2002;
Incyte Genomics, Inc. (US)
FEATURES
Location/Qualifiers
1..2181
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: 7485090CBI"
BASE COUNT 635 a 451 c 419 g 676 t
ORIGIN
Query Match 36.2%; Score 171.8; DB 6; Length 2181;
Best Local Similarity 67.4%; Pred. No. 1.1e-29;
Matches 242; Conservative 0; Mismatches 117; Indels 0; Gaps 0;
Qy 10 TATTCAGTGGCAATTTTCTGTGTATTAATTGGCCGCAATTATCATCATAGTTTTC 69
Dy 1681 TATTCCTGTGGAATTTTCTGTGTATTAATTGGCCGCAATTATCATCATAGTTTTC 1740
```

```

QY 70 TATGGAAGCATGTTTATAGTGTTCATCAAAAGTGCCATAACAGCAACTGAAATACGGAAT 129
    ||| | |||| | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1741 TATATTACTATGTTCTGTTCCATCAAAAACCGCTTCACAGACCACAGAGTAAGGAAT 1800
    || || |||| || |||| || |||| || |||| || |||| || |||| || |||| ||
QY 130 CAAGTTAAAAAGAGATGATCCTTGCCAAACGTTTTCTTTATAGTATTTACTGATGCA 189
    || || |||| || |||| || |||| || |||| || |||| || |||| || |||| ||
Db 1801 TGTTTTGAAGAGAGGTGGCTGCTTGCAAAATCGTTCTTTTATAGTGTCTCTGATGCC 1860
    || || |||| || || |||| || || |||| || || |||| || || |||| || || ||
QY 190 TTATGCTGGATACCCATTTTGTAGCGGAACCTCTTCACTGCTTCAGGTAGTAATACCA 249
    | |||| || || || |||| || || |||| || || |||| || || |||| || || ||
Db 1861 ATCTCTGGATTCCGTATTTAGTTAAATCTTCCCTCTCCGGGTGGAATACCA 1920
    || || || || || || || || || || || || || || || || || || || || ||
QY 250 GGTACCATAACTCTTTGGTAGTGGTTATTCGCCATTACAGTGCCTTGAACCCA 309
    || || || || || || || || || || || || || || || || || || || || ||
Db 1921 GACACAATGACTTCCCTGGATAGTATTTTTCCTCCAGTTAACAGTGTCTTGAATCCA 1980
    || || || || || || || || || || || || || || || || || || || || ||
QY 310 ATTCTCTATCTCTGACCAAGACCAATTAAAGAAATGATTCATCGGTTTGGCATFAA 368
    || || || || || || || || || || || || || || || || || || || || ||
Db 1981 ATCCTCTATCTCTCACACCAACTTTTAAAGGACAAGTTGAACACAGCTGCTGCACAA 2039
    || || || || || || || || || || || || || || || || || || || || ||

```

Search completed: September 25, 2003, 13:58:08  
Job time : 2168 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 25, 2003, 12:09:23 ; Search time 222 Seconds  
(without alignments)  
5763.665 Million cell updates/sec

Title: US-10-049-568-1  
Perfect score: 474  
Sequence: 1 gccacgatttatttcagtgcc.....ggagatgcacactgaggttaa 474

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues  
Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :			
N_Geneseq_19Jun03:*			
1:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:*		
2:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:*		
3:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:*		
4:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:*		
5:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:*		
6:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:*		
7:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:*		
8:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:*		
9:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:*		
10:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:*		
11:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:*		
12:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:*		
13:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:*		
14:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:*		
15:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:*		
16:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:*		
17:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:*		
18:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:*		
19:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:*		
20:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:*		
21:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:*		
22:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:*		
23:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:*		
24:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:*		
25:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/NA2003.DAT:*		

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	474	100.0	474	22 AAF28059	Human HGR101 G-pr
2	439.2	92.7	1089	22 AAI99584	Human expressed po
3	439.2	92.7	1162	23 ABK43573	DNA encoding novel
4	439.2	92.7	1191	22 AAD06507	Human CON222 G pro
5	439.2	92.7	1804	21 AAZ90524	Human GPCR protein
6	439.2	92.7	1804	24 ABX73054	Human G-protein co
7	439.2	92.7	2274	25 ABK42570	Human G protein-co
8	439.2	92.7	2467	20 AAZ25345	Human LGR7 long fo

9	439.2	92.7	3584	20 AAZ25346	Human LGR7 short f
10	438.8	92.6	530	22 AAI99584	Human expressed po
11	438.8	92.6	530	22 AAL35644	Human musculoskele
12	438.8	92.6	530	22 ABA06471	Human cDNA SEQ ID
13	438.8	92.6	530	22 AAS28950	cDNA encoding for
14	438.8	92.6	530	22 AAS29573	Human endocrine po
15	438.8	92.6	530	22 AAS30187	DNA encoding rena
16	438.8	92.6	530	22 AAS34845	cDNA encoding nove
17	438.8	92.6	530	23 ABK43875	DNA encoding novel
18	438.8	92.6	530	24 ABY83808	Human polynucleoti
19	438.8	92.6	530	24 ABT07785	Novel human nuclei
20	438.8	92.6	530	25 ABX58632	cDNA encoding nove
21	437.6	92.3	1126	25 ACA04807	cDNA encoding huma
22	436.6	92.1	612	24 ABX73081	Human GPCR cDNA #4
23	353.2	74.5	420	21 AAC75735	Human ORFX ORF1290
24	171.8	36.2	1065	24 ABL40194	Human G protein-co
25	171.8	36.2	1068	22 AAS07943	Human cDNA encodin
26	171.8	36.2	1473	24 ABL40192	Human G protein-co
27	171.8	36.2	1545	24 ABL40191	Human G protein-co
28	171.8	36.2	1830	24 ABL40188	Human G protein-co
29	171.8	36.2	2142	24 ABK51947	cDNA encoding huma
30	171.8	36.2	2181	24 AAF8580	Human GCRC-1 cDNA
31	171.8	36.2	2190	24 ABL40197	Human G protein-co
32	171.8	36.2	2214	24 ABK51944	cDNA encoding huma
33	171.8	36.2	2262	24 ABL40196	Human G protein-co
34	171.8	36.2	2339	25 AAL51335	Human LGR9 G-prote
35	168.6	35.6	1121	25 ABT33349	NOVX DNA sequence
36	156.8	33.1	321	24 AAD32027	Human novel G-prot
37	154.4	32.6	636	21 AA44932	Human secreted exp
38	107.2	22.6	432	24 AAX73106	Rat variant of hum
39	106.8	22.5	1018	22 AAH51001	Human nGPCR37 codi
40	106.8	22.5	1018	24 ABS70234	DNA encoding human
41	48.8	10.3	1015	22 AAS57085	cDNA encoding Dros
42	48.8	10.3	1015	23 ABL07293	Drosophila melanog
43	44	9.3	4292	21 AAF21792	Human breast and o
44	44	9.3	5644	24 ABR83490	Human cDNA differe
45	40.2	8.5	134499	21 AAF22286	BAC containing rep

# ALIGNMENTS

RESULT 1	
AAAF28059	
ID	AAF28059 standard; cDNA: 474 BP.
XX	
AC	AAF28059;
XX	
DT	23-MAY-2001 (first entry)
DE	Human HGR101 G-protein coupled receptor coding sequence.
XX	
XX	Human; HGR101: G-protein coupled receptor; infection; pain; cancer;
KW	diabetes; obesity; eating disorder; asthma; Parkinson's disease;
KW	hypotension; osteoporosis; myocardial infarction; migraine; allergy;
KW	psychotic disorder; neurological disorder; dyskinesia; vaccine; ss.
XX	
OS	Homo sapiens.
XX	
PH	Key
CDS	Location/Qualifiers
FT	1..474
FT	/*tag= a
FT	/product= "HGR101"
FT	/partial
XX	
PN	WO200114548-A2.
XX	
PD	01-MAR-2001.
XX	
PF	09-AUG-2000; 2000WO-EP07723.
XX	
PR	19-AUG-1999; 99EP-0116345.
XX	

PA (MERE ) MERCK PATENT GMBH.  
 XX Dueker K;  
 XX WPI; 2001-226617/23.  
 DR P-PSDB; AAB35407.  
 XX  
 PT Novel G-coupled protein receptor, HGR101 useful for treating diseases  
 PT such as microbial infections, cancers, obesity, asthma, diabetes,  
 PT hypertension, osteoporosis, myocardial infarction, stroke, ulcer,  
 PT allergy  
 XX  
 PS Claim 5; Page 35-36; 36pp; English.  
 XX  
 CC The present invention provides the protein and coding sequences for a  
 CC novel human G-protein coupled receptor, designated HGR101. The sequences  
 CC are useful in the diagnosis, prevention and treatment of diseases  
 CC including infections, pain, cancer, diabetes, anorexia, bulimia, asthma,  
 CC Parkinson's disease, acute heart failure, hypotension, hypertension,  
 CC urinary retention, osteoporosis, angina pectoris, myocardial infarction,  
 CC stroke, ulcers, allergies, benign prostatic hypertrophy, migraine,  
 CC vomiting, psychotic and neurological disorders such as anxiety,  
 CC schizophrenia, manic depression, depression, delirium, dementia and  
 CC severe mental retardation, and dyskinesias including Huntington's disease  
 CC and Gilles de la Tourette's syndrome. The present sequence is the HGR101  
 CC partial coding sequence.  
 XX  
 SQ Sequence 474 BP; 144 A; 92 C; 88 G; 150 T; 0 other;  
 Query Match 100.0%; Score 474; DB 22; Length 474;  
 Best Local Similarity 100.0%; Pred. No. 4.9e-119; Indels 0; Gaps 0;  
 Matches 474; Conservative 0; Mismatches 0;  
 QY 1 GCCCAGATTATTCAGTGGCAATTTCTTGGTATTATTTGGCGCATTTATCATCA 60  
 Db 1 GCCCAGATTATTCAGTGGCAATTTCTTGGTATTATTTGGCGCATTTATCATCA 60  
 QY 61 GTTTTTCCTATGGAAGCATGTTTATAGTGTTCATCAAGTGCATACAGCACTGAA 120  
 Db 61 GTTTTTCCTATGGAAGCATGTTTATAGTGTTCATCAAGTGCATACAGCACTGAA 120  
 QY 121 ATACGGAATCAAGTAAAAAGAGATGATCCTTGCCAAACGTTTTTCTTTATAGTATT 180  
 Db 121 ATACGGAATCAAGTAAAAAGAGATGATCCTTGCCAAACGTTTTTCTTTATAGTATT 180  
 QY 181 ACTGATCATATGCTGGATACCCATTTTGTAGCGGAACCCCTTTCACCTGTCAGGTA 240  
 Db 181 ACTGATCATATGCTGGATACCCATTTTGTAGCGGAACCCCTTTCACCTGTCAGGTA 240  
 QY 241 GAAATACCAAGTACCAATACCTCTTGGGTAGTGGTATTTCGCCATTACAGTGCT 300  
 Db 241 GAAATACCAAGTACCAATACCTCTTGGGTAGTGGTATTTCGCCATTACAGTGCT 300  
 QY 301 TTGAACCAATCTCTATCTACTGACCAAGACCATTTTAAAGAAATGATTCATCGGTTT 360  
 Db 301 TTGAACCAATCTCTATCTACTGACCAAGACCATTTTAAAGAAATGATTCATCGGTTT 360  
 QY 361 TGGCATACATACAGCAAGAAATCTATGACAGCAAGAGTATCAGAAACATATGCTC 420  
 Db 361 TGGCATACATACAGCAAGAAATCTATGACAGCAAGAGTATCAGAAACATATGCTC 420  
 QY 421 CATCATTCATCTGGGGGAAATGTGGCACTGCGAGGAGATGCCACTGAGTTAA 474  
 Db 421 CATCATTCATCTGGGGGAAATGTGGCACTGCGAGGAGATGCCACTGAGTTAA 474  
 RESULT 2  
 AAI99557  
 ID AAI99557 standard; cDNA; 1089 BP.  
 XX  
 AC AAI99557;  
 XX  
 DT 04-JAN-2002 (first entry)

XX Human expressed polynucleotide SEQ ID NO 20.  
 DE  
 XX  
 KW Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;  
 KW immunosuppressive; anti-inflammatory; anti-HIV; antibacterial; vulnery;  
 KW antiparkinsonian; antiskilling; antianaemic; antiarthritic; cancer;  
 KW antineumatic; hepatotropic; cerebroprotective; antiinflammatory;  
 KW antiallergic; antidiabetic; antilulcer; anticonvulsant; antifungal;  
 KW antiparasitic; cardiant; immune disorder; cardiovascular disorder;  
 KW neurological disease; infection; nephrotropic; gene therapy; vaccine;  
 KW ss.  
 XX  
 XX Homo sapiens.  
 XX WO200155387-A1.  
 PN  
 XX 02-AUG-2001.  
 PD  
 XX 17-JAN-2001; 2001WO-US01310.  
 XX 31-JAN-2000; 2000US-0179065.  
 PR 04-FEB-2000; 2000US-0180628.  
 PR 24-FEB-2000; 2000US-0184664.  
 PR 02-MAR-2000; 2000US-0186350.  
 PR 16-MAR-2000; 2000US-0189874.  
 PR 17-MAR-2000; 2000US-0190076.  
 PR 18-APR-2000; 2000US-0198123.  
 PR 19-MAY-2000; 2000US-0205515.  
 PR 07-JUN-2000; 2000US-0209467.  
 PR 28-JUN-2000; 2000US-0214886.  
 PR 30-JUN-2000; 2000US-0215135.  
 PR 07-JUL-2000; 2000US-0216647.  
 PR 07-JUL-2000; 2000US-0216880.  
 PR 11-JUL-2000; 2000US-0217487.  
 PR 14-JUL-2000; 2000US-0217496.  
 PR 26-JUL-2000; 2000US-0218290.  
 PR 26-JUL-2000; 2000US-0220963.  
 PR 26-JUL-2000; 2000US-0220964.  
 PR 14-AUG-2000; 2000US-0224518.  
 PR 14-AUG-2000; 2000US-0224519.  
 PR 14-AUG-2000; 2000US-0225213.  
 PR 14-AUG-2000; 2000US-0225214.  
 PR 14-AUG-2000; 2000US-0225266.  
 PR 14-AUG-2000; 2000US-0225267.  
 PR 14-AUG-2000; 2000US-0225268.  
 PR 14-AUG-2000; 2000US-0225270.  
 PR 14-AUG-2000; 2000US-0225447.  
 PR 14-AUG-2000; 2000US-0225757.  
 PR 14-AUG-2000; 2000US-0225758.  
 PR 14-AUG-2000; 2000US-0225759.  
 PR 18-AUG-2000; 2000US-0226279.  
 PR 22-AUG-2000; 2000US-0226681.  
 PR 22-AUG-2000; 2000US-0226688.  
 PR 23-AUG-2000; 2000US-0227182.  
 PR 30-AUG-2000; 2000US-0228924.  
 PR 01-SEP-2000; 2000US-0229287.  
 PR 01-SEP-2000; 2000US-0229343.  
 PR 01-SEP-2000; 2000US-0229344.  
 PR 01-SEP-2000; 2000US-0229345.  
 PR 05-SEP-2000; 2000US-0229509.  
 PR 05-SEP-2000; 2000US-0229513.  
 PR 06-SEP-2000; 2000US-0230437.  
 PR 06-SEP-2000; 2000US-0230438.  
 PR 08-SEP-2000; 2000US-0231242.  
 PR 08-SEP-2000; 2000US-0231243.  
 PR 08-SEP-2000; 2000US-0231244.  
 PR 08-SEP-2000; 2000US-0231413.  
 PR 08-SEP-2000; 2000US-0231414.  
 PR 08-SEP-2000; 2000US-0232080.  
 PR 08-SEP-2000; 2000US-0232081.  
 PR 12-SEP-2000; 2000US-0231968.  
 PR 14-SEP-2000; 2000US-0232397.



PR 14-SEP-2000; 2000US-0232398.  
 PR 14-SEP-2000; 2000US-0232399.  
 PR 14-SEP-2000; 2000US-0232400.  
 PR 14-SEP-2000; 2000US-0232401.  
 PR 14-SEP-2000; 2000US-0233063.  
 PR 14-SEP-2000; 2000US-0233064.  
 PR 14-SEP-2000; 2000US-0233065.  
 PR 21-SEP-2000; 2000US-0233223.  
 PR 21-SEP-2000; 2000US-0234274.  
 PR 25-SEP-2000; 2000US-0234997.  
 PR 25-SEP-2000; 2000US-0234998.  
 PR 26-SEP-2000; 2000US-0235484.  
 PR 27-SEP-2000; 2000US-0235834.  
 PR 27-SEP-2000; 2000US-0235836.  
 PR 29-SEP-2000; 2000US-0236327.  
 PR 29-SEP-2000; 2000US-0236367.  
 PR 29-SEP-2000; 2000US-0236368.  
 PR 29-SEP-2000; 2000US-0236369.  
 PR 29-SEP-2000; 2000US-0236370.  
 PR 02-OCT-2000; 2000US-0236802.  
 PR 02-OCT-2000; 2000US-0237037.  
 PR 02-OCT-2000; 2000US-0237038.  
 PR 02-OCT-2000; 2000US-0237039.  
 PR 02-OCT-2000; 2000US-0237040.  
 PR 13-OCT-2000; 2000US-0239935.  
 PR 13-OCT-2000; 2000US-0239937.  
 PR 20-OCT-2000; 2000US-0240960.  
 PR 20-OCT-2000; 2000US-0241221.  
 PR 20-OCT-2000; 2000US-0241785.  
 PR 20-OCT-2000; 2000US-0241786.  
 PR 20-OCT-2000; 2000US-0241787.  
 PR 20-OCT-2000; 2000US-0241808.  
 PR 20-OCT-2000; 2000US-0241809.  
 PR 20-OCT-2000; 2000US-0241826.  
 PR 01-NOV-2000; 2000US-0244617.  
 PR 08-NOV-2000; 2000US-0246474.  
 PR 08-NOV-2000; 2000US-0246475.  
 PR 08-NOV-2000; 2000US-0246476.  
 PR 08-NOV-2000; 2000US-0246477.  
 PR 08-NOV-2000; 2000US-0246478.  
 PR 08-NOV-2000; 2000US-0246523.  
 PR 08-NOV-2000; 2000US-0246524.  
 PR 08-NOV-2000; 2000US-0246525.  
 PR 08-NOV-2000; 2000US-0246526.  
 PR 08-NOV-2000; 2000US-0246527.  
 PR 08-NOV-2000; 2000US-0246528.  
 PR 08-NOV-2000; 2000US-0246532.  
 PR 08-NOV-2000; 2000US-0246609.  
 PR 08-NOV-2000; 2000US-0246610.  
 PR 08-NOV-2000; 2000US-0246611.  
 PR 08-NOV-2000; 2000US-0246613.  
 PR 17-NOV-2000; 2000US-0249207.  
 PR 17-NOV-2000; 2000US-0249208.  
 PR 17-NOV-2000; 2000US-0249209.  
 PR 17-NOV-2000; 2000US-0249210.  
 PR 17-NOV-2000; 2000US-0249211.  
 PR 17-NOV-2000; 2000US-0249212.  
 PR 17-NOV-2000; 2000US-0249213.  
 PR 17-NOV-2000; 2000US-0249214.  
 PR 17-NOV-2000; 2000US-0249215.  
 PR 17-NOV-2000; 2000US-0249216.  
 PR 17-NOV-2000; 2000US-0249217.  
 PR 17-NOV-2000; 2000US-0249218.  
 PR 17-NOV-2000; 2000US-0249244.  
 PR 17-NOV-2000; 2000US-0249245.  
 PR 17-NOV-2000; 2000US-0249264.  
 PR 17-NOV-2000; 2000US-0249265.  
 PR 17-NOV-2000; 2000US-0249267.  
 PR 17-NOV-2000; 2000US-0249297.  
 PR 17-NOV-2000; 2000US-0249299.  
 PR 17-NOV-2000; 2000US-0249300.  
 PR 01-DEC-2000; 2000US-0250160.  
 PR 01-DEC-2000; 2000US-0250391.  
 PR 05-DEC-2000; 2000US-0251030.

PR 05-DEC-2000; 2000US-0251988.  
 PR 05-DEC-2000; 2000US-0256719.  
 PR 06-DEC-2000; 2000US-0251479.  
 PR 08-DEC-2000; 2000US-0251856.  
 PR 08-DEC-2000; 2000US-0251858.  
 PR 08-DEC-2000; 2000US-0251869.  
 PR 08-DEC-2000; 2000US-0251989.  
 PR 08-DEC-2000; 2000US-0251990.  
 PR 11-DEC-2000; 2000US-0254097.  
 PR 05-JAN-2001; 2001US-0259678.  
 XX  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 XX Rosen CA, Barash SC, Ruben SM;  
 XX  
 XX WPI; 2001-465573/50.  
 DR P-PSDB; AAM99945.  
 XX  
 XX Isolated digestive system associated polypeptide for treating,  
 PT preventing and/ or prognosing disorders related to the digestive system  
 PT including digestive system cancers and also for testing and detection  
 PT e.g. diagnosis -  
 XX  
 PS Claim 1; SEQ ID NO 20; 509pp + Sequence Listing; English.  
 XX  
 CC The invention relates to novel genes (AAI99548-AAI99604) and proteins  
 CC (AAM99936-AAM99984) useful for preventing, treating or ameliorating  
 CC medical conditions e.g. by protein or gene therapy. The genes are  
 CC isolated from a range of human tissues disclosed in the specification.  
 CC The nucleic acids, proteins, antibodies and (ant)agonists are useful  
 CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast  
 CC and ovarian cancer and other cancers of the adrenal gland, bone, bone  
 CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;  
 CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune  
 CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's  
 CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative  
 CC colitis; (c) cardiovascular disorders such as myocardial ischaemia;  
 CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and  
 CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal  
 CC and parasitic infections.  
 CC Note: The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 1089 BP; 386 A; 186 C; 186 G; 331 T; 0 other;  
 Query Match 92.7%; Score 439.2; DB 22; Length 1089;  
 Best Local Similarity 96.8%; Pred No. 1.ee-109;  
 Matches 459; Conservative 0; Mismatches 13; Indels 2; Gaps 1;  
 Qy 1 GCCAGATTATTCAGTGGCAATTTTCTGGTATTAATTTGGCGCATTTATCATCATATA 60  
 Db 20 GCCAGATTATTCAGTGGCAATTTTCTGGTATTAATTTGGCGCATTTATCATCATATA 79  
 Qy 61 GTTTTTCCTATGGAAGCATGTTTATAGCTTCATCAAGTGCCTATACAGCAACTGAA 120  
 Db 80 GTTTTTCCTATGGAAGCATGTTTATAGCTTCATCAAGTGCCTATACAGCAACTGAA 139  
 Qy 121 ATACGGATCAAGTTTAAAAAGAGATGATCCTTGCCTTCCAAACGTTTTTCTTTATAGTATTT 180  
 Db 140 ATACGGATCAAGTTTAAAAAGAGATGATCCTTGCCTTCCAAACGTTTTTCTTTATAGTATTT 199  
 Qy 181 ACTGATCATATTATGCTGGATACCCATTTTGTAGCGAAACCTCTTTCATCTCTTCAGGTA 240  
 Db 200 ACTGATCATATTATGCTGGATACCCATTTTGTAGTGAATTTCTTTACTGCTTCAGGTA 259  
 Qy 241 GAATATACAGGTACCATACCTCTTGGGTAGTAGTGGTATTTCTGCCATTAACAGTCT 300  
 Db 260 GAAATACAGGTACCATACCTCTTGGGTAGTAGTGGTATTTCTGCCATTAACAGTCT 319  
 Qy 301 TTGAACCAATTTCTTATCTCTGACCAAGACCATTTTAAAGAAATGATTCATCGGTTT 360  
 Db 320 TTGAACCAATTTCTTATCTCTGACCAAGACCATTTTAAAGAAATGATTCATCGGTTT 379

QY 361 TGGCACTAAGTACAGACAAAGAAATCTATGGACAGCAAGGTATCAGARAACATATGCTC 420  
Db 380 TGGTATTAAGTACAGACAAAGAAATCTATGGACAGCAAGG--TCAGARAACATATGCTC 437  
QY 421 CATCATTCATCTGGGGGAAATGTGGCCACTGCAGGAGATGCCACCTGAGTTAA 474  
Db 438 CATCATTCATCTGGGTGGAAATGTGGCCACTGCAGGAGATGCCACCTGAGTTAA 491

RESULT 3

ABK43573/c  
ID ABK43573 standard; cdna; 1162 BP.

XX AC ABK43573;

XX DT 05-JUN-2002 (first entry)

XX DE DNA encoding novel central nervous system protein #153.

XX KW Central nervous system; CNS; autoimmune disease; rheumatoid arthritis;  
KW hyperproliferative disorder; neoplasm; cardiovascular disorder;  
KW cardiac arrest; cerebrovascular disorder; ischaemia; angiogenesis;  
KW nervous system disorder; Alzheimer's disease; AIDS; ocular disorder;  
KW acquired immunodeficiency virus; dysphagia; gastrointestinal disorder;  
KW adenocarcinoma; reproductive system disorder; testicular feminisation;  
KW endocrine disorder; diabetes; cancer; leukaemia; neovascularisation;  
KW respiratory disorder; renal disorder; kidney failure; blood disorder;  
KW myocardial infarction; wound healing; cell proliferation; skin aging;  
KW food additive; food preservative; gene therapy; gene; ss.

XX OS Homo sapiens.

XX PN WO200155318-A2.

XX PD 02-AUG-2001.

XX PF 17-JAN-2001; 2001WO-US01332.

XX PR 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0215647.  
PR 07-JUL-2000; 2000US-0215880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225277.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226686.  
PR 22-AUG-2000; 2000US-0227182.

PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233085.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246509.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.

PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249287.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
XX  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Barash SC, Ruben SM;  
XX  
XX WPI; 2001-581633/65.  
XX P-PSDB; AAU87243.  
XX  
XX New isolated nucleic acid encoding a protein for diagnosing,  
XX preventing, treating or ameliorating medical conditions and used as  
XX food additives or preservatives -  
XX  
XX Claim 1; SEQ ID No 163; 837pp; English.  
XX  
XX The invention describes an isolated nucleic acid molecule (I) encoding a  
XX novel central nervous system protein. (I) and polypeptides (III) encoded  
XX by (I), are used to treat a medical conditions and in diagnosis of a  
XX pathological condition. Disorders which are diagnosed or treated include  
XX autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative  
XX disorders e.g. neoplasms of the breast or liver, cardiovascular disorders  
XX e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia,  
XX angiodenesis, nervous system disorders e.g. Alzheimer's disease and  
XX amyotrophic lateral sclerosis, infections caused by bacteria, viruses  
XX e.g. Acquired immunodeficiency virus (AIDS) and fungi, ocular disorders  
XX e.g. corneal infection, gastrointestinal disorders e.g. dysphagia,  
XX adenocarcinomas and irritable bowel syndrome, reproductive system  
XX disorders e.g. testicular feminisation, endocrine disorders e.g. diabetes  
XX and pituitary dwarfism, cancers and disorders at the cellular level e.g.  
XX leukaemia, disorders involving neovascularisation e.g. malignancies,  
XX respiratory disorders e.g. nonallergic rhinitis, renal disorders e.g.  
XX acute kidney failure and blood related disorders e.g. myocardial  
XX infarction. The polypeptides can also be used to aid wound healing and  
XX epithelial cell proliferation, to prevent skin aging due to sunburn, to  
XX maintain organs before transplantation, for supporting cell culture of  
XX primary tissues, to regenerate tissues and in chemotaxis. The  
XX polypeptides can also be used as a food additive or preservative to  
XX increase or decrease storage capabilities, fat content, lipid, protein,  
XX  
Query Match 92.7%; Score 439.2; DB 23; Length 1162;  
Best local Similarity 96.8%; Pred. No. 1.8e-109;  
Matches 459; Conservative 0; Mismatches 13; Indels 2; Gaps 1;  
QY 1 GCCACGATTATTCAGTGGCAATTTCTTGGTATTATTTGGCCGCAATTATCATCATTA 60  
|||||

DB 1080 GCCAGATTATTCAGTGGCAATTTCTTGGTATTATTTGGCCGCAATTATCATCATTA 1021  
QY 61 GTTTTTCCTATGGAAGCATGTTTATAGTGTTCATCAAAAGTCCCATACCAACTGAA 120  
|||||  
DB 1020 GTTTTTCCTATGGAAGCATGTTTATAGTGTTCATCAAAAGTCCCATACCAACTGAA 961  
QY 121 ATACGGAAATCAAGTTAAARAGAGATGATCCCTGCCAAGCTTTTCTTTATAGTATT 180  
|||||  
DB 960 ATACGGAAATCAAGTTAAARAGAGATGATCCCTGCCAAGCTTTTCTTTATAGTATT 901  
QY 181 ACTGATGCATTATGCTGGATACCCATTTTTFAGCGAAAGCTTTCACTGCTTCAGGTA 240  
|||||  
DB 900 ACTGATGCATTATGCTGGATACCCATTTTTFAGCGAAATTTCTTCACTGCTTCAGGTA 841  
QY 241 GAAATACAGGTACCATACCTCTGGGTAGTGTGTTATTTCTGCCATTTACAGTCT 300  
|||||  
DB 840 GAAATACAGGTACCATACCTCTGGGTAGTGTGTTATTTCTGCCATTTACAGTCT 781  
QY 301 TTGAACCAATTCCTACTCTGACCAACAGACCATTTAAAGAAATGATTCATCGGTT 360  
|||||  
DB 780 TTGAACCAATTCCTACTCTGACCAACAGACCATTTAAAGAAATGATTCATCGGTT 721  
QY 361 TGGCATAACTACAGACAAAGAAATCTATGGACAGCAAGATATCAGAAAACATATGCTC 420  
|||||  
DB 720 TGGTATAACTACAGACAAAGAAATCTATGGACAGCAAGATATCAGAAAACATATGCTC 663  
QY 421 CATCATTTATCTGGGGGAAATCTGGCCACTGCAGGAGATGCCACCTGAGTTAA 474  
|||||  
DB 662 CATCATTTATCTGGGTGAAATGTGCCACTGCAGGAGATGCCACCTGAGTTAA 609  
RESULT 4  
AAD06507  
ID AAD06507 standard; DNA; 1191 Bp.  
XX  
XX AC AAD06507;  
XX  
XX DT 10-AUG-2001 (first entry)  
XX  
XX DE Human CON222 G protein-coupled receptor DNA.  
XX  
XX KW Human; G protein-coupled receptor; GPCR; CON222 protein; schizophrenia;  
KW neuroleptic; nootropic; neuroprotective; bipolar disease; psychotropic;  
KW neurological disorder; psychiatric disease; neurosis; anxiety; neuritis;  
KW attention deficit hyperactivity disorder; neuroasthenia; senile dementia;  
KW affective disorder; neuropathy; Alzheimer's disease; Parkinson's disease;  
KW depression; migraine; genetic screening; ds.  
XX  
XX OS Homo sapiens.  
XX  
XX FH Key Location/Qualifiers  
XX CDS 1..1191  
XX FT /\*tag= a  
XX FT /product= "Human G protein-coupled receptor protein"  
XX  
XX WO200131014-A2.  
XX  
XX PD 03-MAY-2001.  
XX  
XX PF 27-OCT-2000; 2000WO-US29601.  
XX  
XX PR 27-OCT-1999; 99US-0427653.  
XX PR 27-OCT-1999; 99US-0427859.  
XX PR 27-OCT-1999; 99US-0428020.  
XX PR 27-OCT-1999; 99US-0428114.  
XX PR 28-OCT-1999; 99US-0429517.  
XX PR 28-OCT-1999; 99US-0429555.  
XX PR 28-OCT-1999; 99US-0429676.  
XX PR 28-OCT-1999; 99US-0429695.  
XX PR 03-DEC-1999; 99US-0454399.  
XX PR 12-JAN-2000; 2000US-0481794.  
XX  
XX PA (PHAA ) PHARMACIA & UPJOHN CO.

XX Vogeli G, Wood LS, Merchant K;  
 XX WPI; 2001-328653/34.  
 XX P-PSDB; AAE02498.  
 XX Seven transmembrane receptor polypeptides and polynucleotides, useful  
 XX for treating neurological or psychiatric disorders, e.g. schizophrenia,  
 XX as well as for identifying compounds useful for treating schizophrenia  
 XX  
 XX Claim 22; Page 15-16; 215pp; English.  
 XX The invention relates to human G protein-coupled receptor (GPCR) and  
 XX their corresponding DNA molecules. GPCR is also referred as seven  
 XX transmembrane receptor. G protein-coupled receptor protein is useful for  
 XX treating neurological disorder, particularly schizophrenia. GPCR protein  
 XX is also useful for identifying compounds useful for treating  
 XX schizophrenia. These compounds are also useful for treating other  
 XX neurological and psychiatric diseases, e.g. depression, anxiety, bipolar  
 XX disease, affective disorders, attention deficit hyperactivity disorder/  
 XX attention deficit disorder, epilepsy, neuritis, neuroathenia, neuroathy,  
 XX neurosis, Alzheimer's disease, Parkinson's disease, migraine and senile  
 XX dementia. The invention also provides genetic screening procedures that  
 XX entail analysing a person's genome with respect to GPCR. The vectors are  
 XX useful for the recombinant production of the GPCR's. The present DNA  
 XX sequence encodes human CON222 G protein-coupled receptor (GPCR) protein.  
 XX  
 XX Sequence 1191 BP; 340 A; 229 C; 226 G; 396 T; 0 other;  
 XX  
 XX Query Match 92.7%; Score 439.2; DB 22; Length 1191;  
 XX Best Local Similarity 96.8%; Pred. No. 1.8e-109;  
 XX Matches 459; Conservative 0; Mismatches 13; Indels 2; Gaps 1;  
 XX  
 XX 1 GCCAGATATTTCAGTGGCAATTTCTTGTTATTAATTTGGCGCATTTATCATCATA 60  
 XX  
 XX 643 GCCAGATATTTCAGTGGCAATTTCTTGTTATTAATTTGGCGCATTTATCATCATA 702  
 XX  
 XX 61 GTTTTTCCATGAAGCAATGTTTATAGTCTTCATCAAGTGGCCATACAGCAACTGAA 120  
 XX  
 XX 703 GTTTTTCCATGAAGCAATGTTTATAGTCTTCATCAAGTGGCCATACAGCAACTGAA 762  
 XX  
 XX 121 ATACGGATCAAGTAAAGAGATGATCTTCCCAAGCGTTTCTTTATAGTATTT 180  
 XX  
 XX 763 ATACGGATCAAGTAAAGAGATGATCTTCCCAAGCGTTTCTTTATAGTATTT 822  
 XX  
 XX 181 ACTGATGCATTATGCTGGATACCCATTTTCTAGCGAAACCTCTTTCACGCTTCAGGTA 240  
 XX  
 XX 823 ACTGATGCATTATGCTGGATACCCATTTTCTAGCGAAATCTTTCACGCTTCAGGTA 882  
 XX  
 XX 241 GAAATACAGGTACCATTAACCTCTTGGGTAGTATGTTTCTTGCATTAACAGTGT 300  
 XX  
 XX 883 GAAATACAGGTACCATTAACCTCTTGGGTAGTATGTTTCTTGCATTAACAGTGT 942  
 XX  
 XX 301 TTGAACCCCAATCTCTATCTCTGACCAAGACCATTTAAAGAAATGATTCATCGGTTT 360  
 XX  
 XX 943 TTGAACCCCAATCTCTATCTCTGACCAAGACCATTTAAAGAAATGATTCATCGGTTT 1002  
 XX  
 XX 361 TGGCATACTACAGCAAGAAATCTTATGACAGCAAGGATATCAGAAACATATGCTC 420  
 XX  
 XX 1003 TGGTATACTACAGCAAGAAATCTTATGACAGCAAGG--TCAGAAACATATGCTC 1060  
 XX  
 XX 421 CATCATTCATCTGGGGGAATGTTGGCCACTGACGAGAGTGCACCTGAGTTAA 474  
 XX  
 XX 1061 CATCATTCATCTGGGGGAATGTTGGCCACTGACGAGAGTGCACCTGAGTTAA 1114  
 XX  
 XX  
 XX RESULT 5  
 XX ID AAZ90524 standard; cDNA; 1804 BP.  
 XX AC AAZ90524;  
 XX

DT 05-JUN-2000 (first entry)  
 XX Human GPCR protein (HGPRP) encoding cDNA (clone ID 2488822).  
 DE Human; G protein coupled protein receptor; HGPRP; cell proliferation;  
 XX neurological; immune disorder; cytostatic; anti-arteriosclerotic;  
 KW anti-atherosclerotic; hepatotropic; antiinflammatory; virucide; leukemia;  
 KW immunomodulatory; anemia; asthma; gastrointestinal; anti-epileptic;  
 KW anti-Alzheimer's; anti-Parkinsonian; gene therapy; ss.  
 XX Homo sapiens.  
 OS  
 XX WO200015793-A2.  
 PN  
 XX 23-MAR-2000.  
 PD  
 XX 17-SEP-1999; 99WO-US20958.  
 PF  
 XX 17-SEP-1998; 98US-0156513.  
 PR  
 XX (INCY-) INCYTE PHARM INC.  
 PA  
 XX Bandman O, Lal P, Tang YT, Corley NC, Guegler KJ, Gorgone GA;  
 PI Baughn MR;  
 PI WPI; 2000-271432/23.  
 DR P-PSDB; AAY57286.  
 DR  
 XX Human G protein coupled protein receptor peptides useful for the  
 PT prevention, diagnosis and treatment of cell proliferative, neurological  
 PT and immune disorders -  
 PT  
 XX Claim 9; Page 69-70; 71pp; English.  
 PS  
 XX The invention provides human G protein coupled protein receptor (HGPRP)  
 CC polypeptides and polynucleotides encoding them. The polypeptides can be  
 CC produced by standard recombinant methodology. The polynucleotides and  
 CC polypeptides may be used in the prevention, treatment and diagnosis of  
 CC diseases associated with their inappropriate expression. Diseases that  
 CC can be treated are cell proliferative disorders (e.g. arteriosclerosis,  
 CC atherosclerosis and hepatitis), cancers (e.g. leukemia, melanomas and  
 CC adenocarcinoma), immune disorders (e.g. anemia, asthma and Crohn's  
 CC disease) and neurological disorders (e.g. epilepsy, Alzheimer's disease  
 CC and Parkinson's disease). The anti-HGPRP antibodies may also be used as  
 CC diagnostic agents for detecting the presence of HGPRP polypeptides in  
 CC samples (e.g. by enzyme linked immunosorbant assay (ELISA)). Sequences  
 CC AAZ90521-526 represent cDNA fragments encoding the HGPRP polypeptides  
 CC (AAZ57283-288).  
 CC  
 XX Sequence 1804 BP; 563 A; 318 C; 327 G; 596 T; 0 other;  
 XX  
 XX Query Match 92.7%; Score 439.2; DB 21; Length 1804;  
 XX Best Local Similarity 96.8%; Pred. No. 2.1e-109;  
 XX Matches 459; Conservative 0; Mismatches 13; Indels 2; Gaps 1;  
 XX  
 XX 1 GCCAGATTTATTCAGTGGCAATTTCTTGTTATTAATTTGGCGCATTTATCATCATA 60  
 XX  
 XX 792 GCCAGATTTATTCAGTGGCAATTTCTTGTTATTAATTTGGCGCATTTATCATCATA 851  
 XX  
 XX 61 GTTTTTTCCCTATGGAAGCATGTTTATAGTGTTCATCAAGTGCCTAACAGCAACTGAA 120  
 XX  
 XX 852 GTTTTTTCCCTATGGAAGCATGTTTATAGTGTTCATCAAGTGCCTAACAGCAACTGAA 911  
 XX  
 XX 121 ATACGGATCAAGTAAAGAGATGATCTTGCCTCAACAGTCTTTTCTTTATAGTATTT 180  
 XX  
 XX 912 ATACGGATCAAGTAAAGAGATGATCTTGCCTCAACAGTCTTTTCTTTATAGTATTT 971  
 XX  
 XX 181 ACTGATGCATTATGCTGGATACCCATTTTGTAGTGCCTAACAGCAACTGAA 240  
 XX  
 XX 972 ACTGATGCATTATGCTGGATACCCATTTTGTAGTGCCTAACAGCAACTGAA 1031  
 XX  
 XX 241 GAAATACAGGTACCATTAACCTCTTGGGTAGTATGTTTCTTGCATTAACAGTGT 300  
 XX  
 XX

Db 1032 GAAATACAGGTACATTAACCTCTTGGGTAGTGATTTTATCTGCCCCATTAACAGTCT 1091  
 QY 301 TTGAACCAATCTCTATCTGACCAACAGACCATTTAAAGAAATGATTCATCGGTTT 360  
 Db 1092 TTGAACCAATCTCTATCTGACCAACAGACCATTTAAAGAAATGATTCATCGGTTT 1151  
 QY 361 TGCATTAATCTACAGCAAAAGAAATCTATGGACAGCAAGGTATCAGAAACATATGCTC 420  
 Db 1152 TGTATTAATCTACAGCAAAAGAAATCTATGGACAGCAAGGTATCAGAAACATATGCTC 1209  
 QY 421 CATCATTAATCTGCGGGGAAATGTGGCCACTGCAGAGATGCCACCTGAGTTAA 474  
 Db 1210 CATCATTAATCTGCGGGGAAATGTGGCCACTGCAGAGATGCCACCTGAGTTAA 1263

RESULT 6  
 ID ABX73054 standard; cDNA; 1804 BP.  
 XX AC ABX73054;  
 XX DT 14-MAR-2003 (first entry)  
 XX DE Human G-protein coupled receptor cDNA #4, Incyte clone 2488822CB1.  
 XX KW Human; ss; gene; G-protein coupled receptor; GPCR; neoplastic disorder;  
 KW neurological disorder; immune disorder; cytostatic; pancreatic cancer;  
 KW follicular carcinoma of the thyroid; leiomyoma of the uterus; epilepsy;  
 KW interstitial nephritis; cancer.  
 XX OS Homo sapiens.  
 XX PN US2002106655-A1.  
 XX PD 08-AUG-2002.  
 XX PF 28-JUN-2001; 2001US-0895686.  
 XX PR 17-SEP-1998; 98US-0156513.  
 XX PA (BAND/) BANDMAN O.  
 PA (LALP/) LAL P G.  
 PA (TANG/) TANG Y T.  
 PA (BAUG/) BAUGHN M R.  
 XX PI Bandman O, Lal PG, Tang YT, Baughn MR;  
 XX WPI: 2002-697866/75.  
 XX P-PSDB: ABU54868.

New cDNAs encoding G protein coupled receptors are useful for the diagnosis, prognosis, treatment and evaluation of therapies for neoplastic, neurological and immune disorders -  
 Claim 2; Page 31-32; 61pp; English.  
 The invention relates to an isolated cDNA encoding G-protein coupled receptor (GPCR) appearing as ABU54865-ABU54870.  
 Also included are fragments of the cDNAs, species variants having at least 75% identity to the cDNAs, vectors comprising the cDNAs, a host cells comprising the above vectors, producing a protein (comprising the protein), using a cDNA to detect expression of a nucleic acid in a sample or to screen for compounds or molecules which bind to the cDNAs, using the GPCR proteins to screen compounds or molecules for ligands, using a GPCR protein to prepare and purify antibodies, an anti-GPCR antibody and using the antibody to detect expression of a GPCR protein in a sample and is diagnostic of cancer.  
 The invention is useful for the diagnosis, prognosis, treatment and evaluation of therapies for neoplastic, neurological and immune disorders, particularly follicular carcinoma of the thyroid, leiomyoma of the uterus, pancreatic cancer, epilepsy, interstitial nephritis and immune response as a complication of cancer. The present sequence

CC encodes a human GPCR protein of the invention.  
 XX SQ Sequence 1804 BP; 563 A; 318 C; 327 G; 596 T; 0 other;  
 Query Match 92.7%; Score 439.2; DB 24; Length 1804;  
 Best Local Similarity 96.8%; Pred. No. 2.1e-109;  
 Matches 459; Conservative 0; Mismatches 13; Indels 2; Gaps 1;  
 QY 1 GCCAGATTTATTCAGTGGCAATTTTCTGGTATTAATTTGGCGCATTTATCATCATA 50  
 Db 792 GCCAGATTTATTCAGTGGCAATTTTCTGGTATTAATTTGGCGCATTTATCATCATA 851  
 QY 61 GTTTTTCCTATGGAAGCATGTTTATAGTGTTCATCAAGTGCATACAGCACTGAA 120  
 Db 852 GTTTTTCCTATGGAAGCATGTTTATAGTGTTCATCAAGTGCATACAGCACTGAA 911  
 QY 121 ATACGGATCAAGTTTAAAGAGAGATGATCCTTGGCCAAACGTTTTTCTTTATAGTATTT 180  
 Db 912 ATACGGATCAAGTTTAAAGAGAGATGATCCTTGGCCAAACGTTTTTCTTTATAGTATTT 971  
 QY 181 ACTGATGATATGCTGGATACCCATTTTGGTAGTGGTTATTTCTGCCATTAACAGTCT 240  
 Db 972 ACTGATGATATGCTGGATACCCATTTTGGTAGTGGTTATTTCTGCCATTTACAGTGA 1031  
 QY 241 GAAATACAGGTACCAATACCTCTTGGGTAGTGGTTATTTCTGCCATTAACAGTCT 300  
 Db 1032 GAAATACAGGTACCAATACCTCTTGGGTAGTGGTTATTTCTGCCATTTACAGTGTCT 1091  
 QY 301 TTGAACCAATCTCTATCTGACCAACAGACCATTTAAAGAAATGATTCATCGGTTT 360  
 Db 1092 TTGAACCAATCTCTATCTGACCAACAGACCATTTAAAGAAATGATTCATCGGTTT 1151  
 QY 361 TGSCATTAATACAGCAAAAGAAATCTATGGACAGCAAGGTATTCAGAAACATATGCTC 420  
 Db 1152 TGTATTAATCTACAGCAAAAGAAATCTATGGACAGCAAGGTATTCAGAAACATATGCTC 1209  
 QY 421 CATCATTAATCTGCGGGGAAATGTGGCCACTGCAGAGATGCCACCTGAGTTAA 474  
 Db 1210 CATCATTAATCTGCGGGGAAATGTGGCCACTGCAGAGATGCCACCTGAGTTAA 1263

RESULT 7  
 ID ABZ42570 standard; DNA; 2274 BP.  
 XX AC ABZ42570;  
 XX DT 04-MAR-2003 (first entry)  
 XX DE Human G protein-coupled receptor LGR7 nucleotide SEQ ID NO:622.  
 KW G protein-coupled receptor; GPCR; antigenic peptide; gene therapy;  
 KW protein-coupled receptor modulator; antibody; immune-related disease;  
 KW growth-related disease; cell regenerative-related disease; AIDS; cancer;  
 KW immunological-related cell proliferative disease; autoimmune disease;  
 KW Alzheimer's disease; atherosclerosis; infection; osteoarthritis; allergy;  
 KW osteoporosis; cardiomyopathy; inflammation; Crohn's disease; diabetes;  
 KW graft versus host disease; Parkinson's disease; multiple sclerosis; pain;  
 KW psoriasis; anxiety; depression; schizophrenia; dementia; memory loss;  
 KW mental retardation; epilepsy; asthma; tuberculosis; obesity; nausea;  
 KW hypertension; hypotension; renal disorder; rheumatoid arthritis; trauma;  
 KW ulcer; gene; ds.  
 XX OS Homo sapiens.  
 XX PN WO200261087-A2.  
 XX PD 08-AUG-2002.  
 XX PF 19-DEC-2001; 2001WO-US50107.  
 XX PR 19-DEC-2000; 2000US-257144P.

PA (LIFE-) LIFESPAN BIOSCIENCES INC.  
 XX Burner GC, Roush CL, Brown JP;  
 XX WPI; 2003-046718/04.  
 DR P-PSDB; ABP81724.  
 XX  
 XX New isolated antigenic peptides e.g., for G protein-coupled receptors  
 PT (GPCR), useful for diagnosing and designing drugs for treating  
 PT conditions in which GPCRs are involved, e.g. AIDS, Alzheimer's disease,  
 PT cancer or autoimmune diseases -  
 XX  
 XX Disclosure; Fig 1; 523pp; English.  
 XX  
 XX The present invention describes antigenic peptides (I) comprising:  
 CC (a) any one of 1601 sequences (see ABP82019 to ABP83619) of 12-24 amino  
 CC acids. Also described: (1) an assay for the detection of a particular  
 CC G protein-coupled receptor (GPCR) or a candidate polypeptide in a sample;  
 CC and (2) an isolated antibody having high specificity and high affinity  
 CC or avidity for a particular GPCR. (1) can be used as GPCR modulators and  
 CC in gene therapy. The antigenic peptides for GPCRs are useful in detecting  
 CC an antibody against a particular GPCR, and in the production of specific  
 CC antibodies. The peptides and antibodies are also useful for detecting the  
 CC presence or absence of corresponding GPCRs. The antigenic peptides for  
 CC GPCRs and antibodies are useful for diagnosing and designing drugs for  
 CC treating immune-related diseases, growth-related diseases, cell  
 CC regeneration-related disease, immunological-related cell proliferative  
 CC diseases, or autoimmune diseases, e.g. AIDS, Alzheimer's disease,  
 CC atherosclerosis, bacterial, fungal, protozoan or viral infections,  
 CC osteoarthritis, osteoporosis, cancer, cardiomyopathy, chronic and acute  
 CC inflammation, allergies, Crohn's disease, diabetes, graft versus host  
 CC disease, Parkinson's disease, multiple sclerosis, pain, psoriasis,  
 CC anxiety, depression, schizophrenia, dementia, mental retardation, memory  
 CC loss, epilepsy, asthma, tuberculosis, obesity, nausea, hypertension,  
 CC hypotension, renal disorders, rheumatoid arthritis, trauma, ulcers, or  
 CC any other disorder in which GPCRs are involved. The antibodies may be  
 CC used in immunoassays and immunodiagnosis. ABP42523 to ABP43869 encode  
 CC GPCR proteins given in ABP81675 to ABP82018, which are used in the  
 CC exemplification of the present invention.  
 XX  
 XX Sequence 2274 BP; 669 A; 460 C; 421 G; 724 T; 0 other;

Query Match 92.7%; Score 439.2; DB 25; Length 2274;  
 Best Local Similarity 96.8%; Pred. No. 2.3e-109;  
 Matches 459; Conservative 0; Mismatches 13; Indels 2; Gaps 1;  
 QY 1 GCCAGATTATTCAGTGGCAATTTCTTGGTATTAATTTGGCGCATTTATCATCA 60  
 DB 1726 GCCAGATTATTCAGTGGCAATTTCTTGGTATTAATTTGGCGCATTTATCATCA 1785  
 QY 61 GTTTTTCCTATGGAGCATGTTTATAGTGTTCNCAAGTGCCTAACGCACTGAA 120  
 DB 1786 GTTTTTCCTATGGAGCATGTTTATAGTGTTCNCAAGTGCCTAACGCACTGAA 1845  
 QY 121 ATACGGATCAAGTATTAAGAGATGATCTTCCCAACGTTTCTTTATAGTATTT 180  
 DB 1846 ATACGGATCAAGTATTAAGAGATGATCTTCCCAACGTTTCTTTATAGTATTT 1905  
 QY 181 ACTGATGATATGCTGGATACCAATTTTGTAGCAGCAACCTCTTCTCAGTCA 240  
 DB 1906 ACTGATGATATGCTGGATACCAATTTTGTAGTGAATTTCTTCTCAGTCA 1965  
 QY 241 GAAATACAGTACCAATACCTCTTGGGTAGTATGTTTCTTGGCATTAACAGTCT 300  
 DB 1966 GAAATACAGTACCAATACCTCTTGGGTAGTATGTTTCTTGGCATTAACAGTCT 2025  
 QY 301 TTGAACCAATCTCTACTACTGACCAAGACCAATTTAAAGAAATGATTCATCGGTT 360  
 DB 2026 TTGAACCAATCTCTACTACTGACCAAGACCAATTTAAAGAAATGATTCATCGGTT 2085  
 QY 361 TGGCAATACACAGCAAAATCTATGACAGCAAGGATATCAGAAACATATGCTC 420  
 DB 2086 TGGTAACTACAGCAAAATCTATGACAGCAAGGATATCAGAAACATATGCTC 2143

QY 421 CATCATTCATCTGGGGGAAATCTGCCACTGCCAGAGATGCCACCTGAGTTAA 474  
 DB 2144 CATCATTCATCTGGGGGAAATCTGCCACTGCCAGAGATGCCACCTGAGTTAA 2197  
 RESULT 8  
 ID AA225345 standard; cDNA; 2467 BP.  
 XX AA225345;  
 XX 20-DEC-1999 (first entry)  
 XX Human LGR7 long form nucleotide sequence.  
 DE Human LGR7 long form nucleotide sequence.  
 XX Human; LGR4; LGR5; LGR7; G-protein coupled receptor; gene therapy;  
 KW extracellular leucine rich repeat region; mapping; identification; ss.  
 XX Homo sapiens.  
 OS Homo sapiens.  
 XX WO9948921-A1.  
 PN 30-SEP-1999.  
 PD 25-MAR-1999; 99WO-US06573.  
 PF 26-MAR-1998; 98US-0079501.  
 PR (STRD ) UNIV LELAND STANFORD JUNIOR.  
 PA (ORGA ) ORGANON NV.  
 XX Hsueh AJW, Hsu SY, Liang S, Van Der Spek PJ;  
 PI WPI; 1999-591074/50.  
 DR P-PSDB; AAY42170.  
 XX New G-protein coupled receptors, useful for identifying their own  
 PT ligands -  
 XX Claim 4; Fig 3; 54pp; English.  
 XX The present sequence encodes the human G-protein coupled receptor  
 CC having extracellular leucine rich repeat regions, designated LGR7 long  
 CC form. The LGR4, LGR5 and LGR7 proteins are used to identify ligands for  
 CC the receptor. The polypeptides and/or polynucleotides are also useful  
 CC for homologous or related genes, producing compositions that modulate  
 CC the expression or function of the receptors, gene therapy, mapping  
 CC functional regions of the receptors, studying associated physiological  
 CC pathways, in vivo prophylactic and therapeutic purposes, as immunogens  
 CC for producing antibodies, and for identifying biologically active  
 CC agents. The polypeptides contain a G-protein coupled seven  
 CC transmembrane region and a leucine rich repeat extracellular domain.  
 CC These regions capture and facilitate optimal orientation of its ligand.  
 CC The proteins are also expressed in diverse tissues.  
 XX Sequence 2467 BP; 747 A; 487 C; 474 G; 759 T; 0 other;  
 Query Match 92.7%; Score 439.2; DB 20; Length 2467;  
 Best Local Similarity 96.8%; Pred. No. 2.3e-109;  
 Matches 459; Conservative 0; Mismatches 13; Indels 2; Gaps 1;  
 QY 1 GCCAGATTATTCAGTGGCAATTTCTTGGTATTAATTTGGCGCATTTATCATCA 60  
 DB 1919 GCCAGATTATTCAGTGGCAATTTCTTGGTATTAATTTGGCGCATTTATCATCA 1978  
 QY 61 GTTTTTCCTATGGAGCATGTTTATAGTGTTCNCAAGTGCCTAACGCACTGAA 120  
 DB 1979 GTTTTTCCTATGGAGCATGTTTATAGTGTTCNCAAGTGCCTAACGCACTGAA 2038  
 QY 121 ATACGGATCAAGTATTAAGAGATGATCTTCCCAACGTTTCTTTATAGTATTT 180  
 DB 2039 ATACGGATCAAGTATTAAGAGATGATCTTCCCAACGTTTCTTTATAGTATTT 2098

QY 181 ACTGATGATATGCTGGATACCACTTTTGTAGGAGAAACCTTTTCACTGCTTCAGGTA 240  
 Db 2099 ACTGATGATATGCTGGATACCACTTTTGTAGGAGAAATTTTCACTGCTTCAGGTA 2158  
 QY 241 GAATACAGGATACCACTGCTGGTGTAGTGTATTTGCTGCTTACAGTGT 300  
 Db 2159 GAATACAGGATACCACTGCTGGTGTAGTGTATTTGCTGCTTACAGTGT 2218  
 QY 301 TTGAACCAATTTCTTACTCTGACCAAGACCAATTTTAAAGAAATGATTCATCGGTT 360  
 Db 2219 TTGAACCAATTTCTTACTCTGACCAAGACCAATTTTAAAGAAATGATTCATCGGTT 2278  
 QY 361 TGGCAATACACAGCAAGAAATCTATGACAGCAAGATGATCAGAAACATATGCTC 420  
 Db 2279 TGGTATACACAGCAAGAAATCTATGACAGCAAGATGATCAGAAACATATGCTC 2336  
 QY 421 CATCATCTCTGGGGGAAATGTGGCCACTGACAGAGATGCCACCTGAGTTAA 474  
 Db 2337 CATCATCTCTGGGGGAAATGTGGCCACTGACAGAGATGCCACCTGAGTTAA 2390

RESULT 9  
 AAZ25346  
 ID AAZ25346 standard; cDNA; 3584 BP.  
 AC AAZ25346;  
 XX  
 XX  
 XX 20-DEC-1999 (first entry)  
 XX Human LGR7 short form nucleotide sequence.  
 XX Human; LGR4; LGR5; LGR7; G-protein coupled receptor; gene therapy;  
 KW extracellular leucine rich repeat region; mapping; identification; ss.  
 KW  
 OS Homo sapiens.  
 XX  
 XX WO948921-A1.  
 XX 30-SEP-1999.  
 XX 25-MAR-1999; 99WO-US06573.  
 XX 26-MAR-1998; 98US-0079501.  
 XX (STRD ) UNIV LELAND STANFORD JUNIOR.  
 XX (ORGA ) ORGANON NV.  
 XX Hsueh AJW, Hsu SY, Liang S, Van Der Spek PJ;  
 XX WPI; 1999-591074/50.  
 XX P-PSDB; AAY42171.  
 XX New G-protein coupled receptors, useful for identifying their own  
 PT ligands -  
 XX  
 XX Claim 4; Fig 4; 54pp; English.

The present sequence encodes the human G-protein coupled receptor  
 CC having extracellular leucine rich repeat regions, designated LGR7 short  
 CC form. The LGR4, LGR5 and LGR7 proteins are used to identify ligands for  
 CC the receptor. The polypeptides and/or polynucleotides are also useful  
 CC for homologous or related genes, producing compositions that modulate  
 CC the expression or function of the receptors, gene therapy, mapping  
 CC functional regions of the receptors, studying associated physiological  
 CC pathways, in vivo prophylactic and therapeutic purposes, as immunogens  
 CC for producing antibodies, and for identifying biologically active  
 CC agents. The polypeptides contain a G-protein coupled seven  
 CC transmembrane region and a leucine rich repeat extracellular domain.  
 CC These regions capture and facilitate optimal orientation of its ligand.  
 CC The proteins are also expressed in diverse tissues.  
 XX  
 XX Sequence 3584 BP; 1124 A; 670 C; 647 G; 1139 T; 4 other;

Query Match 92.7%; Score 439.2; DB 20; Length 3584;  
 Best Local Similarity 96.8%; Pred. No. 2.6e-109;  
 Matches 459; Conservative 0; Mismatches 13; Indels 2; Gaps 1;  
 QY 1 GCCCAGATTTATTCAGTGGCAATTTTCTTGGTATTAATTTGGCCCATTTATCATATA 60  
 Db 1744 GCCCAGATTTATTCAGTGGCAATTTTCTTGGTATTAATTTGGCCCATTTATCATATA 1803  
 QY 61 GTTTTTTCTTATGGAACCATGTTTTATAGTGTTCATCAAAAGTGCCTAATAACAGCACTGAA 120  
 Db 1804 GTTTTTTCTTATGGAACCATGTTTTATAGTGTTCATCAAAAGTGCCTAATAACAGCACTGAA 1863  
 QY 121 ATACGGAATCAAGTTAAATAAGAGATGATCCTTGGCCAAACGTTTTTCTTATAGTATTT 180  
 Db 1864 ATACGGAATCAAGTTAAATAAGAGATGATCCTTGGCCAAACGTTTTTCTTATAGTATTT 1923  
 QY 181 ACTGATGATATGCTGATACCACTTTTGTAGGAAACCTCTTTCACTGCTTCAGGTA 240  
 Db 1924 ACTGATGATATGCTGATACCACTTTTGTAGGAAATTTTCTTCACTGCTTCAGGTA 1983  
 QY 241 GAAATACAGGTACCATTAACCTCTTGGGATGATGTTATTTCTGCCATTAACAGTGT 300  
 Db 1984 GAAATACAGGTACCATTAACCTCTTGGGATGATGTTATTTCTGCCATTAACAGTGT 2043  
 QY 301 TTGAACCAATTTCTTACTCTGACCAAGACCAATTTTAAAGAAATGATTCATCGGTT 360  
 Db 2044 TTGAACCAATTTCTTACTCTGACCAAGACCAATTTTAAAGAAATGATTCATCGGTT 2103  
 QY 361 TGGCAATACACAGCAAGAAATCTATGACAGCAAGATGATCAGAAACATATGCTC 420  
 Db 2104 TGGTATACACAGCAAGAAATCTATGACAGCAAGATGATCAGAAACATATGCTC 2161  
 QY 421 CATCATCTCTGGGGGAAATGTGGCCACTGACAGAGATGCCACCTGAGTTAA 474  
 Db 2162 CATCATCTCTGGGGGAAATGTGGCCACTGACAGAGATGCCACCTGAGTTAA 2215

RESULT 10  
 AAI99584  
 ID AAI99584 standard; cDNA; 530 BP.  
 XX AC AAI99584;  
 XX  
 XX 04-JAN-2002 (first entry)  
 XX Human expressed polynucleotide SEQ ID NO 47.

Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;  
 KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;  
 KW antiparkinsonian; antisickling; antianemic; antithratic; cancer;  
 KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;  
 KW antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;  
 KW antiparasitic; cardiant; immune disorder; cardiovascular disorder;  
 KW neurological disease; infection; nephrotropic; gene therapy; vaccine;  
 XX ss.

XX Homo sapiens.  
 XX OS  
 XX WO200155387-A1.  
 XX  
 XX 02-AUG-2001.  
 XX  
 XX 17-JAN-2001; 2001WO-US01310.  
 XX  
 XX 31-JAN-2000; 2000US-0179065.  
 XX 04-FEB-2000; 2000US-0180628.  
 XX 24-FEB-2000; 2000US-0184664.  
 XX 02-MAR-2000; 2000US-0186350.  
 XX 16-MAR-2000; 2000US-0189874.  
 XX 17-MAR-2000; 2000US-0190076.  
 XX 18-APR-2000; 2000US-0198123.  
 XX 19-MAY-2000; 2000US-0205515.

PR 07-JUN-2000; 2000US-0209467.  
 PR 28-JUN-2000; 2000US-0214886.  
 PR 30-JUN-2000; 2000US-0215135.  
 PR 07-JUL-2000; 2000US-0216647.  
 PR 07-JUL-2000; 2000US-0216880.  
 PR 11-JUL-2000; 2000US-0217487.  
 PR 11-JUL-2000; 2000US-0217496.  
 PR 14-JUL-2000; 2000US-0218290.  
 PR 26-JUL-2000; 2000US-0220963.  
 PR 26-JUL-2000; 2000US-0220964.  
 PR 14-AUG-2000; 2000US-0224518.  
 PR 14-AUG-2000; 2000US-0224519.  
 PR 14-AUG-2000; 2000US-0225213.  
 PR 14-AUG-2000; 2000US-0225214.  
 PR 14-AUG-2000; 2000US-0225266.  
 PR 14-AUG-2000; 2000US-0225267.  
 PR 14-AUG-2000; 2000US-0225268.  
 PR 14-AUG-2000; 2000US-0225270.  
 PR 14-AUG-2000; 2000US-0225271.  
 PR 14-AUG-2000; 2000US-0225275.  
 PR 14-AUG-2000; 2000US-0225278.  
 PR 14-AUG-2000; 2000US-0225279.  
 PR 18-AUG-2000; 2000US-0226279.  
 PR 22-AUG-2000; 2000US-0226681.  
 PR 22-AUG-2000; 2000US-0226686.  
 PR 22-AUG-2000; 2000US-0227182.  
 PR 23-AUG-2000; 2000US-0227009.  
 PR 30-AUG-2000; 2000US-0228924.  
 PR 01-SEP-2000; 2000US-0229287.  
 PR 01-SEP-2000; 2000US-0229343.  
 PR 01-SEP-2000; 2000US-0229344.  
 PR 01-SEP-2000; 2000US-0229345.  
 PR 05-SEP-2000; 2000US-0229509.  
 PR 05-SEP-2000; 2000US-0229513.  
 PR 06-SEP-2000; 2000US-0230437.  
 PR 06-SEP-2000; 2000US-0230438.  
 PR 08-SEP-2000; 2000US-0231242.  
 PR 08-SEP-2000; 2000US-0231243.  
 PR 08-SEP-2000; 2000US-0231244.  
 PR 08-SEP-2000; 2000US-0231413.  
 PR 08-SEP-2000; 2000US-0231414.  
 PR 08-SEP-2000; 2000US-0232080.  
 PR 08-SEP-2000; 2000US-0232081.  
 PR 12-SEP-2000; 2000US-0231268.  
 PR 14-SEP-2000; 2000US-0232397.  
 PR 14-SEP-2000; 2000US-0232398.  
 PR 14-SEP-2000; 2000US-0232399.  
 PR 14-SEP-2000; 2000US-0232400.  
 PR 14-SEP-2000; 2000US-0232401.  
 PR 14-SEP-2000; 2000US-0233063.  
 PR 14-SEP-2000; 2000US-0233064.  
 PR 14-SEP-2000; 2000US-0233065.  
 PR 21-SEP-2000; 2000US-0234223.  
 PR 21-SEP-2000; 2000US-0234274.  
 PR 25-SEP-2000; 2000US-0234997.  
 PR 25-SEP-2000; 2000US-0234998.  
 PR 26-SEP-2000; 2000US-0235484.  
 PR 27-SEP-2000; 2000US-0235694.  
 PR 27-SEP-2000; 2000US-0235836.  
 PR 29-SEP-2000; 2000US-0236327.  
 PR 29-SEP-2000; 2000US-0236367.  
 PR 29-SEP-2000; 2000US-0236368.  
 PR 29-SEP-2000; 2000US-0236369.  
 PR 29-SEP-2000; 2000US-0236370.  
 PR 02-OCT-2000; 2000US-0236802.  
 PR 02-OCT-2000; 2000US-0237037.  
 PR 02-OCT-2000; 2000US-0237038.  
 PR 02-OCT-2000; 2000US-0237039.  
 PR 02-OCT-2000; 2000US-0237040.  
 PR 13-OCT-2000; 2000US-0239935.  
 PR 13-OCT-2000; 2000US-0239937.  
 PR 20-OCT-2000; 2000US-0240960.  
 PR 20-OCT-2000; 2000US-0241221.  
 PR 20-OCT-2000; 2000US-0241785.  
 PR 20-OCT-2000; 2000US-0241786.  
 PR 20-OCT-2000; 2000US-0241787.  
 PR 20-OCT-2000; 2000US-0241808.  
 PR 20-OCT-2000; 2000US-0241809.  
 PR 20-OCT-2000; 2000US-0241826.  
 PR 01-NOV-2000; 2000US-0244617.  
 PR 08-NOV-2000; 2000US-0244674.  
 PR 08-NOV-2000; 2000US-0244675.  
 PR 08-NOV-2000; 2000US-0244676.  
 PR 08-NOV-2000; 2000US-0244677.  
 PR 08-NOV-2000; 2000US-0244678.  
 PR 08-NOV-2000; 2000US-0245233.  
 PR 08-NOV-2000; 2000US-0245234.  
 PR 08-NOV-2000; 2000US-0245252.  
 PR 08-NOV-2000; 2000US-0245256.  
 PR 08-NOV-2000; 2000US-0245257.  
 PR 08-NOV-2000; 2000US-0245287.  
 PR 08-NOV-2000; 2000US-0245288.  
 PR 08-NOV-2000; 2000US-0245322.  
 PR 08-NOV-2000; 2000US-0246609.  
 PR 08-NOV-2000; 2000US-0246610.  
 PR 08-NOV-2000; 2000US-0246611.  
 PR 08-NOV-2000; 2000US-0246613.  
 PR 17-NOV-2000; 2000US-0249207.  
 PR 17-NOV-2000; 2000US-0249208.  
 PR 17-NOV-2000; 2000US-0249209.  
 PR 17-NOV-2000; 2000US-0249210.  
 PR 17-NOV-2000; 2000US-0249211.  
 PR 17-NOV-2000; 2000US-0249212.  
 PR 17-NOV-2000; 2000US-0249213.  
 PR 17-NOV-2000; 2000US-0249214.  
 PR 17-NOV-2000; 2000US-0249215.  
 PR 17-NOV-2000; 2000US-0249216.  
 PR 17-NOV-2000; 2000US-0249217.  
 PR 17-NOV-2000; 2000US-0249218.  
 PR 17-NOV-2000; 2000US-0249244.  
 PR 17-NOV-2000; 2000US-0249245.  
 PR 17-NOV-2000; 2000US-0249264.  
 PR 17-NOV-2000; 2000US-0249265.  
 PR 17-NOV-2000; 2000US-0249297.  
 PR 17-NOV-2000; 2000US-0249299.  
 PR 17-NOV-2000; 2000US-0249300.  
 PR 01-DEC-2000; 2000US-0250160.  
 PR 01-DEC-2000; 2000US-0250391.  
 PR 05-DEC-2000; 2000US-0251030.  
 PR 05-DEC-2000; 2000US-0251988.  
 PR 05-DEC-2000; 2000US-0256719.  
 PR 06-DEC-2000; 2000US-0251479.  
 PR 08-DEC-2000; 2000US-0251856.  
 PR 08-DEC-2000; 2000US-0251868.  
 PR 08-DEC-2000; 2000US-0251869.  
 PR 08-DEC-2000; 2000US-0251989.  
 PR 08-DEC-2000; 2000US-0251990.  
 PR 11-DEC-2000; 2000US-0254097.  
 PR 05-JAN-2001; 2001US-0259678.  
 (HUMA-) HUMAN GENOME SCI INC.  
 PI Rosen CA, Barash SC, Ruben SM;  
 WPI; 2001-465573/50.  
 P-PSDB; AAM99972.  
 PT Isolated digestive system associated polypeptide for treating,  
 preventing and/ or prognosing disorders related to the digestive system  
 including digestive system cancers and also for testing and detection  
 e.g. diagnosis -  
 XX Claim 1; SEQ ID NO 47; 509pp + Sequence Listing; English.  
 PA The invention relates to novel genes (AAI99548-AAI99604) and proteins  
 (AAM99836-AAM99984) useful for preventing, treating or ameliorating  
 CC medical conditions e.g. by protein or gene therapy. The genes are



CC isolated from a range of human tissues disclosed in the specification.  
 CC The nucleic acids, proteins, antibodies and (ant)agonists are useful  
 CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast  
 CC and ovarian cancer and other cancers of the adrenal gland, bone, bone  
 CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;  
 CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune  
 CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's  
 CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative  
 CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;  
 CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and  
 CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal  
 CC and parasitic infections.  
 CC Note: The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at [ftp.wipo.int/pub/published\\_pct\\_sequences](http://ftp.wipo.int/pub/published_pct_sequences).

XX SQ Sequence 530 BP; 158 A; 105 C; 95 G; 171 T; 1 other;

Query Match 92.6%; Score 438.8; DB 22; Length 530;  
 Best Local Similarity 96.6%; Pred. No. 1.9e-109;  
 Matches 458; Conservative 1; Mismatches 13; Indels 2; Gaps 1;  
 QY 1 GCCAGATTATTCAGTGGCAATTTCTTGTTATTAATTTGGCGCAATTCATCATATA 60  
 Db 2 GCCAGATTATTCAGTGGCAATTTCTTGTTATTAATTTGGCGCAATTCATCATATA 61  
 QY 61 GTTTTTCCTATGGAGCATGTTTATAGTGTTCATCAAGTGGCCATACGCACTGAA 120  
 Db 62 GTTTTTCCTATGGAGCATGTTTATAGTGTTCATCAAGTGGCCATACGCACTGAA 121  
 QY 121 ATACGGAATCAAGTTAAAGAGAGATGATCCTTGCCAAACGTTTTCCTTTATAGTATT 180  
 Db 122 ATACGGAATCAAGTTAAAGAGAGATGATCCTTGCCAAACGTTTTCCTTTATAGTATT 181  
 QY 181 ACTGATGCATTTAGTGGATACCATTTTGTAGCGAAACCTCTTTCACGTGTCAGGTA 240  
 Db 182 ACTGATGCATTTAGTGGATACCATTTTGTAGCGAAACCTCTTTCACGTGTCAGGTA 241  
 QY 241 GAATACCAAGTACCATACCTCTGGGTAGTGTGTTATTCGTCATTAACAGTCT 300  
 Db 242 GAATACCAAGTACCATACCTCTGGGTAGTGTGTTATTCGTCATTAACAGTCT 301  
 QY 301 TTGACCAATTTCTTACTCTGACCAACAGACCATTTAAAGAAATGATTCATCGGTT 360  
 Db 302 TTGACCAATTTCTTACTCTGACCAACAGACCATTTAAAGAAATGATTCATCGGTT 361  
 QY 361 TGSCATACTACAGACAAAGAAATCTATGGACAGCAAGGATACAGAAACATATGCTC 420  
 Db 362 TGGTATACTACAGACAAAGAAATCTATGGACAGCAAGGATACAGAAACATATGCTC 419  
 QY 421 CATCATTCATCTGGGGGGAATGCGCACTGCGAGAGATGCCACCTGAGTTAA 474  
 Db 420 CATCATTCATCTGGGTGGAATGCGCACTGCGAGAGATGCCACCTGAGTTAA 473

RESULT 11  
 AAL35644

ID AAL35644 standard; cDNA; 530 BP.

XX AC AAL35644;

XX DT 08-JAN-2002 (first entry)

XX DE Human musculoskeletal system related polynucleotide SEQ ID NO 988.

XX KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;  
 KW antiallergic; hepatocytic; antidiabetic; antiinflammatory; antiulcer;  
 KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;  
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;  
 KW neurological disease; infection; human; secreted protein;  
 KW musculoskeletal system; ss.  
 XX OS Homo sapiens.

XX WO200155367-A1.  
 PN 02-AUG-2001.  
 XX 17-JAN-2001; 2001WO-US01338.  
 XX 31-JAN-2000; 2000US-0179065.  
 PR 04-FEB-2000; 2000US-0180628.  
 PR 24-FEB-2000; 2000US-0184664.  
 PR 02-MAR-2000; 2000US-0186350.  
 PR 16-MAR-2000; 2000US-0189874.  
 PR 17-MAR-2000; 2000US-0190076.  
 PR 18-APR-2000; 2000US-0198123.  
 PR 19-MAY-2000; 2000US-0205515.  
 PR 07-JUN-2000; 2000US-0209467.  
 PR 28-JUN-2000; 2000US-0214886.  
 PR 30-JUN-2000; 2000US-0215135.  
 PR 07-JUL-2000; 2000US-0216647.  
 PR 11-JUL-2000; 2000US-0217487.  
 PR 11-JUL-2000; 2000US-0217496.  
 PR 14-JUL-2000; 2000US-0218290.  
 PR 26-JUL-2000; 2000US-0220963.  
 PR 26-JUL-2000; 2000US-0220964.  
 PR 14-AUG-2000; 2000US-0224518.  
 PR 14-AUG-2000; 2000US-0224519.  
 PR 14-AUG-2000; 2000US-0225213.  
 PR 14-AUG-2000; 2000US-0225214.  
 PR 14-AUG-2000; 2000US-0225266.  
 PR 14-AUG-2000; 2000US-0225267.  
 PR 14-AUG-2000; 2000US-0225268.  
 PR 14-AUG-2000; 2000US-0225270.  
 PR 14-AUG-2000; 2000US-0225447.  
 PR 14-AUG-2000; 2000US-0225757.  
 PR 14-AUG-2000; 2000US-0225758.  
 PR 14-AUG-2000; 2000US-0225759.  
 PR 18-AUG-2000; 2000US-0226279.  
 PR 22-AUG-2000; 2000US-0226681.  
 PR 22-AUG-2000; 2000US-0226888.  
 PR 22-AUG-2000; 2000US-0227182.  
 PR 23-AUG-2000; 2000US-0227182.  
 PR 30-AUG-2000; 2000US-0228924.  
 PR 01-SEP-2000; 2000US-0229287.  
 PR 01-SEP-2000; 2000US-0229343.  
 PR 01-SEP-2000; 2000US-0229344.  
 PR 01-SEP-2000; 2000US-0229345.  
 PR 05-SEP-2000; 2000US-0229509.  
 PR 05-SEP-2000; 2000US-0229513.  
 PR 06-SEP-2000; 2000US-0230437.  
 PR 06-SEP-2000; 2000US-0230438.  
 PR 08-SEP-2000; 2000US-0231242.  
 PR 08-SEP-2000; 2000US-0231243.  
 PR 08-SEP-2000; 2000US-0231244.  
 PR 08-SEP-2000; 2000US-0231413.  
 PR 08-SEP-2000; 2000US-0231414.  
 PR 08-SEP-2000; 2000US-0232080.  
 PR 08-SEP-2000; 2000US-0232081.  
 PR 12-SEP-2000; 2000US-0231968.  
 PR 14-SEP-2000; 2000US-0232397.  
 PR 14-SEP-2000; 2000US-0232398.  
 PR 14-SEP-2000; 2000US-0232399.  
 PR 14-SEP-2000; 2000US-0232400.  
 PR 14-SEP-2000; 2000US-0232401.  
 PR 14-SEP-2000; 2000US-0233063.  
 PR 14-SEP-2000; 2000US-0233064.  
 PR 14-SEP-2000; 2000US-0233065.  
 PR 21-SEP-2000; 2000US-0234223.  
 PR 21-SEP-2000; 2000US-0234274.  
 PR 25-SEP-2000; 2000US-0234997.  
 PR 25-SEP-2000; 2000US-0234998.  
 PR 26-SEP-2000; 2000US-0235484.  
 PR 27-SEP-2000; 2000US-0235834.



ID ABA06471 standard; cDNA; 530 BP.  
 XX AC ABA06471;  
 XX DT 10-JAN-2002 (first entry)  
 XX DE Human cDNA SEQ ID NO: 137.  
 XX KW Human; gene therapy; neural disorder; immune system disorder;  
 KW muscular disorder; reproductive disorder; gastrointestinal disorder;  
 KW pulmonary disorder; cardiovascular disorder; renal disorder;  
 KW proliferative disorder; inflammation; ss.  
 XX OS Homo sapiens.  
 XX PN WO200154474-A2.  
 XX PD 02-AUG-2001.  
 XX PF 17-JAN-2001; 2001WO-US01349.  
 XX PR 31-JAN-2000; 2000US-179065P.  
 PR 04-FEB-2000; 2000US-180628P.  
 PR 24-FEB-2000; 2000US-184664P.  
 PR 02-MAR-2000; 2000US-186350P.  
 PR 16-MAR-2000; 2000US-189874P.  
 PR 17-MAR-2000; 2000US-190076P.  
 PR 18-APR-2000; 2000US-198123P.  
 PR 19-MAY-2000; 2000US-205515P.  
 PR 07-JUN-2000; 2000US-209467P.  
 PR 28-JUN-2000; 2000US-214885P.  
 PR 30-JUN-2000; 2000US-215135P.  
 PR 07-JUL-2000; 2000US-216647P.  
 PR 07-JUL-2000; 2000US-216880P.  
 PR 11-JUL-2000; 2000US-217487P.  
 PR 11-JUL-2000; 2000US-217496P.  
 PR 14-JUL-2000; 2000US-218290P.  
 PR 26-JUL-2000; 2000US-220963P.  
 PR 26-JUL-2000; 2000US-220964P.  
 PR 14-AUG-2000; 2000US-224518P.  
 PR 14-AUG-2000; 2000US-224519P.  
 PR 14-AUG-2000; 2000US-225213P.  
 PR 14-AUG-2000; 2000US-225214P.  
 PR 14-AUG-2000; 2000US-225266P.  
 PR 14-AUG-2000; 2000US-225267P.  
 PR 14-AUG-2000; 2000US-225268P.  
 PR 14-AUG-2000; 2000US-225270P.  
 PR 14-AUG-2000; 2000US-225447P.  
 PR 14-AUG-2000; 2000US-225757P.  
 PR 14-AUG-2000; 2000US-225758P.  
 PR 14-AUG-2000; 2000US-225759P.  
 PR 18-AUG-2000; 2000US-226279P.  
 PR 22-AUG-2000; 2000US-226681P.  
 PR 22-AUG-2000; 2000US-226688P.  
 PR 22-AUG-2000; 2000US-227182P.  
 PR 23-AUG-2000; 2000US-227009P.  
 PR 30-AUG-2000; 2000US-228924P.  
 PR 01-SEP-2000; 2000US-229287P.  
 PR 01-SEP-2000; 2000US-229343P.  
 PR 01-SEP-2000; 2000US-229344P.  
 PR 01-SEP-2000; 2000US-229345P.  
 PR 05-SEP-2000; 2000US-229509P.  
 PR 05-SEP-2000; 2000US-229513P.  
 PR 06-SEP-2000; 2000US-230437P.  
 PR 06-SEP-2000; 2000US-230438P.  
 PR 08-SEP-2000; 2000US-231242P.  
 PR 08-SEP-2000; 2000US-231243P.  
 PR 08-SEP-2000; 2000US-231244P.  
 PR 08-SEP-2000; 2000US-231413P.  
 PR 08-SEP-2000; 2000US-231414P.  
 PR 08-SEP-2000; 2000US-232080P.  
 PR 08-SEP-2000; 2000US-232081P.  
 PR 12-SEP-2000; 2000US-231965P.  
 PR 14-SEP-2000; 2000US-232397P.  
 PR 14-SEP-2000; 2000US-232398P.  
 PR 14-SEP-2000; 2000US-232399P.  
 PR 14-SEP-2000; 2000US-232400P.  
 PR 14-SEP-2000; 2000US-232401P.  
 PR 14-SEP-2000; 2000US-233063P.  
 PR 14-SEP-2000; 2000US-233064P.  
 PR 14-SEP-2000; 2000US-233065P.  
 PR 21-SEP-2000; 2000US-234223P.  
 PR 21-SEP-2000; 2000US-234274P.  
 PR 25-SEP-2000; 2000US-234997P.  
 PR 25-SEP-2000; 2000US-234998P.  
 PR 26-SEP-2000; 2000US-235484P.  
 PR 27-SEP-2000; 2000US-235834P.  
 PR 27-SEP-2000; 2000US-235835P.  
 PR 29-SEP-2000; 2000US-236327P.  
 PR 29-SEP-2000; 2000US-236367P.  
 PR 29-SEP-2000; 2000US-236368P.  
 PR 29-SEP-2000; 2000US-236369P.  
 PR 29-SEP-2000; 2000US-236370P.  
 PR 02-OCT-2000; 2000US-236802P.  
 PR 02-OCT-2000; 2000US-237037P.  
 PR 02-OCT-2000; 2000US-237038P.  
 PR 02-OCT-2000; 2000US-237039P.  
 PR 02-OCT-2000; 2000US-237040P.  
 PR 13-OCT-2000; 2000US-239935P.  
 PR 13-OCT-2000; 2000US-239937P.  
 PR 20-OCT-2000; 2000US-240960P.  
 PR 20-OCT-2000; 2000US-241221P.  
 PR 20-OCT-2000; 2000US-241785P.  
 PR 20-OCT-2000; 2000US-241786P.  
 PR 20-OCT-2000; 2000US-241787P.  
 PR 20-OCT-2000; 2000US-241808P.  
 PR 20-OCT-2000; 2000US-241809P.  
 PR 01-NOV-2000; 2000US-241826P.  
 PR 08-NOV-2000; 2000US-246474P.  
 PR 08-NOV-2000; 2000US-246475P.  
 PR 08-NOV-2000; 2000US-246476P.  
 PR 08-NOV-2000; 2000US-246477P.  
 PR 08-NOV-2000; 2000US-246478P.  
 PR 08-NOV-2000; 2000US-246523P.  
 PR 08-NOV-2000; 2000US-246524P.  
 PR 08-NOV-2000; 2000US-246525P.  
 PR 08-NOV-2000; 2000US-246526P.  
 PR 08-NOV-2000; 2000US-246527P.  
 PR 08-NOV-2000; 2000US-246528P.  
 PR 08-NOV-2000; 2000US-246532P.  
 PR 08-NOV-2000; 2000US-246609P.  
 PR 08-NOV-2000; 2000US-246610P.  
 PR 08-NOV-2000; 2000US-246611P.  
 PR 17-NOV-2000; 2000US-249207P.  
 PR 17-NOV-2000; 2000US-249208P.  
 PR 17-NOV-2000; 2000US-249209P.  
 PR 17-NOV-2000; 2000US-249210P.  
 PR 17-NOV-2000; 2000US-249211P.  
 PR 17-NOV-2000; 2000US-249212P.  
 PR 17-NOV-2000; 2000US-249213P.  
 PR 17-NOV-2000; 2000US-249214P.  
 PR 17-NOV-2000; 2000US-249215P.  
 PR 17-NOV-2000; 2000US-249216P.  
 PR 17-NOV-2000; 2000US-249217P.  
 PR 17-NOV-2000; 2000US-249218P.  
 PR 17-NOV-2000; 2000US-249244P.  
 PR 17-NOV-2000; 2000US-249245P.  
 PR 17-NOV-2000; 2000US-249264P.  
 PR 17-NOV-2000; 2000US-249265P.  
 PR 17-NOV-2000; 2000US-249297P.  
 PR 17-NOV-2000; 2000US-249299P.  
 PR 17-NOV-2000; 2000US-249300P.  
 PR 01-DEC-2000; 2000US-250160P.  
 PR 01-DEC-2000; 2000US-250391P.

```

PR 05-DEC-2000; 2000US-251030P.
PR 05-DEC-2000; 2000US-251988P.
PR 05-DEC-2000; 2000US-256719P.
PR 06-DEC-2000; 2000US-251472P.
PR 06-DEC-2000; 2000US-251856P.
PR 08-DEC-2000; 2000US-251868P.
PR 08-DEC-2000; 2000US-251869P.
PR 08-DEC-2000; 2000US-251989P.
PR 08-DEC-2000; 2000US-251990P.
PR 11-DEC-2000; 2000US-254097P.
PR 05-JAN-2001; 2001US-259678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-476161/51.
XX P-PSDB; ABB10249.
XX
XX Isolated nucleic acid molecule encoding an inflammation-associated
XX polypeptide is used in preventing, treating or ameliorating a medical
XX condition
XX
XX Claim 1; SEQ ID NO: 137; 859pp + Sequence Listing; English.
XX
XX The present invention provides human cDNAs, proteins and related genomic
XX DNAs. These can be used in the treatment of neural, immune system,
XX muscular, reproductive, gastrointestinal, pulmonary, cardiovascular,
XX renal and proliferative disorders and inflammation. The present sequence
XX is a cDNA of the invention.
XX
XX Sequence 530 BP; 158 A; 105 C; 95 G; 171 T; 1 other;
XX
XX Query Match 92.6%; Score 438.8; DB 22; Length 530;
XX Best Local Similarity 96.6%; Pred. No. 1.9e-109;
XX Matches 458; Conservative 1; Mismatches 13; Indels 2; Gaps 1;
XX
QY 1 GCCAGATTATTCAGTGGCAATTTTCTGGTATTAATTTGGCGCATTTATCATCAPA 60
Db |||||
QY 2 GCCAGATTATTCAGTGGCAATTTTCTGGTATTAATTTGGCGCATTTATCATCAPA 61
Db |||||
QY 61 GTTTTTCCTATGGAAGCATGTTTATAGTTCATCAAGTGGCCATACAGCAACTGAA 120
Db |||||
QY 62 GTTTTTCCTATGGAAGCATGTTTATAGTTCATCAAGTGGCCATACAGCAACTGAA 121
Db |||||
QY 121 ATACGGAATCAAGTAAAGAGAGATGATCCTTGCCAAACGTTTCTTTATAGTATT 180
Db |||||
QY 122 ATACGGAATCAAGTAAAGAGAGATGATCCTTGCCAAACGTTTCTTTATAGTATT 181
Db |||||
QY 181 ACTGATGCAATATGCTGNTACCCATTTTGTAGCGAACCCTTTCACCTGCTCAGSTA 240
Db |||||
QY 182 ACTGATGCAATATGCTGNTACCCATTTTGTAGCGAACCCTTTCACCTGCTCAGSTA 241
Db |||||
QY 241 GAAATACCAAGTACCATAACCTCTTGGGTAGTATGTTTCTGTCATTAACAGTGCT 300
Db |||||
QY 242 GAAATACCAAGTACCATAACCTCTTGGGTAGTATGTTTCTGTCATTAACAGTGCT 301
Db |||||
QY 301 TTGACCCCAATCTCTATACCTCTGACCAAGACCCATTTAAGAAATGATCATCGGTTT 360
Db |||||
QY 302 TTGACCCCAATCTCTATACCTCTGACCAAGACCCATTTAAGAAATGATCATCGGTTT 361
Db |||||
QY 361 TGGCAATACCTACAGCAAGAAATCTATGGACAGCAAGGATTCAGAAACATATGCTC 420
Db |||||
QY 362 TGGTATACCTACAGCAAGAAATCTATGGACAGCAAGGATTCAGAAACATATGCTC 419
Db |||||
QY 421 CATCATTCATCTGGGGGAATGTGGCCACTGCAGGAGATGCCACCTGAGTTAA 474
Db |||||
QY 420 CATCATTCATCTGGGGGAATGTGGCCACTGCAGGAGATGCCACCTGAGTTAA 473
Db |||||

RESULT 13
AAS28950
ID AAS28950 standard; cDNA; 530 BP.

```

```

XX
AC AAS28950;
XX
DT 21-NOV-2001 (first entry)
XX
DE cDNA encoding for human uterine motility-association polypeptide #15.
XX
KW Human; uterine motility-association disorder; uterus; pregnancy;
KW labour; menstrual cycle; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN WO200155201-A1.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01317.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226888.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.

```

PR 14-SEP-2000; 2000US-0232400.  
 PR 14-SEP-2000; 2000US-0232401.  
 PR 14-SEP-2000; 2000US-0233063.  
 PR 14-SEP-2000; 2000US-0233064.  
 PR 14-SEP-2000; 2000US-0233065.  
 PR 21-SEP-2000; 2000US-0234223.  
 PR 21-SEP-2000; 2000US-0234274.  
 PR 25-SEP-2000; 2000US-0234997.  
 PR 25-SEP-2000; 2000US-0234998.  
 PR 26-SEP-2000; 2000US-0235484.  
 PR 27-SEP-2000; 2000US-0235834.  
 PR 27-SEP-2000; 2000US-0235836.  
 PR 29-SEP-2000; 2000US-0236327.  
 PR 29-SEP-2000; 2000US-0236367.  
 PR 29-SEP-2000; 2000US-0236368.  
 PR 29-SEP-2000; 2000US-0236369.  
 PR 29-SEP-2000; 2000US-0236370.  
 PR 02-OCT-2000; 2000US-0236802.  
 PR 02-OCT-2000; 2000US-0237037.  
 PR 02-OCT-2000; 2000US-0237038.  
 PR 02-OCT-2000; 2000US-0237039.  
 PR 02-OCT-2000; 2000US-0237040.  
 PR 13-OCT-2000; 2000US-0239935.  
 PR 13-OCT-2000; 2000US-0239937.  
 PR 20-OCT-2000; 2000US-0240960.  
 PR 20-OCT-2000; 2000US-0241221.  
 PR 20-OCT-2000; 2000US-0241785.  
 PR 20-OCT-2000; 2000US-0241786.  
 PR 20-OCT-2000; 2000US-0241787.  
 PR 20-OCT-2000; 2000US-0241808.  
 PR 20-OCT-2000; 2000US-0241809.  
 PR 20-OCT-2000; 2000US-0241826.  
 PR 01-NOV-2000; 2000US-0244617.  
 PR 08-NOV-2000; 2000US-0244674.  
 PR 08-NOV-2000; 2000US-024475.  
 PR 08-NOV-2000; 2000US-0246476.  
 PR 08-NOV-2000; 2000US-0246477.  
 PR 08-NOV-2000; 2000US-0246478.  
 PR 08-NOV-2000; 2000US-0246523.  
 PR 08-NOV-2000; 2000US-0246524.  
 PR 08-NOV-2000; 2000US-0246525.  
 PR 08-NOV-2000; 2000US-0246526.  
 PR 08-NOV-2000; 2000US-0246527.  
 PR 08-NOV-2000; 2000US-0246528.  
 PR 08-NOV-2000; 2000US-0246532.  
 PR 08-NOV-2000; 2000US-0246609.  
 PR 08-NOV-2000; 2000US-0246610.  
 PR 08-NOV-2000; 2000US-0246611.  
 PR 08-NOV-2000; 2000US-0246613.  
 PR 17-NOV-2000; 2000US-0249207.  
 PR 17-NOV-2000; 2000US-0249208.  
 PR 17-NOV-2000; 2000US-0249209.  
 PR 17-NOV-2000; 2000US-0249210.  
 PR 17-NOV-2000; 2000US-0249211.  
 PR 17-NOV-2000; 2000US-0249212.  
 PR 17-NOV-2000; 2000US-0249213.  
 PR 17-NOV-2000; 2000US-0249214.  
 PR 17-NOV-2000; 2000US-0249215.  
 PR 17-NOV-2000; 2000US-0249216.  
 PR 17-NOV-2000; 2000US-0249217.  
 PR 17-NOV-2000; 2000US-0249218.  
 PR 17-NOV-2000; 2000US-0249244.  
 PR 17-NOV-2000; 2000US-0249245.  
 PR 17-NOV-2000; 2000US-0249264.  
 PR 17-NOV-2000; 2000US-0249265.  
 PR 17-NOV-2000; 2000US-0249297.  
 PR 17-NOV-2000; 2000US-0249299.  
 PR 17-NOV-2000; 2000US-0249300.  
 PR 01-DEC-2000; 2000US-0250160.  
 PR 01-DEC-2000; 2000US-0250391.  
 PR 05-DEC-2000; 2000US-0251030.  
 PR 05-DEC-2000; 2000US-0251988.  
 PR 05-DEC-2000; 2000US-0256719.

PR 06-DEC-2000; 2000US-0251479.  
 PR 08-DEC-2000; 2000US-0251856.  
 PR 08-DEC-2000; 2000US-0251868.  
 PR 08-DEC-2000; 2000US-0251869.  
 PR 08-DEC-2000; 2000US-0251989.  
 PR 08-DEC-2000; 2000US-0251990.  
 PR 11-DEC-2000; 2000US-0254097.  
 PR 05-JAN-2001; 2001US-0259678.  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 XX Rosen CA, Barash SC, Ruben SM;  
 PI WPI; 2001-488777/53.  
 XX P-PSDB; AAU18108.  
 DR Isolated polypeptide and nucleic acid molecules for treating,  
 XX preventing and/or prognosing disorders related to uterine motility  
 XX e.g. disorders associated with pregnancy and the menstrual cycle -  
 XX  
 PS Claim 4; SEQ ID No 25; 524pp; English.  
 XX  
 CC The present invention relates to the isolation of novel human  
 CC uterine motility-association polypeptides (AAU18094-AAU18152),  
 CC and cDNA and genomic sequences encoding for these polypeptides.  
 CC The sequences of the invention are useful in the diagnosis,  
 CC treatment, prevention and/or prognosis of diseases associated  
 CC with uterine motility such as pregnancy and labour, and menstrual  
 CC disorders. The polynucleotide sequences of the invention are also  
 CC useful in gene therapy. AAS28936-AAS28994 represent cDNA sequences  
 CC encoding for novel human uterine motility-association polypeptides.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 530 BP; 158 A; 105 C; 95 G; 171 T; 1 other:  
 Query Match 92.6%; Score 438.8; DB 22; Length 530;  
 Best Local Similarity 96.6%; Pred. No. 1.9e-109;  
 Matches 458; Conservative 1; Mismatches 13; Indels 2; Gaps 1;  
 QY 1 GCCAGATTATTCAGTGGCAATTTTCTTGTTATTAATTTGGCCGATTATCATCATATA 60  
 Db 2 GCCAGATTATTCAGTGGCAATTTTCTTGTTATTAATTTGGCCGATTATCATCATATA 61  
 QY 61 GTTTTTCCTATGAGCAGTCTTTTATAGTGTCTATCAAGTGCATACACCACTGAA 120  
 Db 62 GTTTTTCCTATGAGCAGTCTTTTATAGTGTCTATCAAGTGCATACACCACTGAA 121  
 QY 121 ATACGGAATCAAGTTAAAAAGAGATGATCCTTGCCAAACGTTTTTCTTTATAGTATT 180  
 Db 122 ATACGGAATCAAGTTAAAAAGAGATGATCCTTGCCAAACGTTTTTCTTTATAGTATT 181  
 QY 181 ACTGATGCATTATGCTGGATACCCATTTTGTAGCGAAACCTTTTCTCTTATAGTATT 240  
 Db 182 ACTGATGCATTATGCTGGATACCCATTTTGTAGCGAAATTTTCTCTTATAGTATT 241  
 QY 241 GAAATACCAAGTACCAATCAACCTTTGGTAGTGTGTTATCTGCCATTACAGTCT 300  
 Db 242 GAAATACCAAGTACCAATCAACCTTTGGTAGTGTGTTATCTGCCATTACAGTCT 301  
 QY 301 TTGAACCCAAATTCCTATCTACTCTGACCAACAGCATTATAAGAAATGATCATCGGTTT 360  
 Db 302 TTGAACCCAAATTCCTATCTACTCTGACCAACAGCATTATAAGAAATGATCATCGGTTT 361  
 QY 361 TGGCATACTACAGACAAAGAAATCTATGAGACACAGATATCAGAAACATATGCTC 420  
 Db 362 TGGTATACTACAGACAAAGAAATCTATGAGACACAGATATCAGAAACATATGCTC 420  
 QY 421 CATCATTCATCTGGGGGAATGTGGCCACTGCGAGATGCCACCTGAGTTAA 474  
 Db 420 CATCATTCATCTGGGGGAATGTGGCCACTGCGAGATGCCACCTGAGTTAA 473

RESULT 14  
AAS29573  
ID AAS29573 standard; cDNA; 530 BP.  
XX AC AAS29573;  
XX DT 21-NOV-2001 (first entry)  
XX DE Human endocrine polypeptide encoding cDNA SEQ ID No 73.  
XX KW Endocrine protein; human; mouse; rabbit; goat; horse; food additive;  
KW cat; dog; chicken; sheep; immunosuppressive; antiarthritic; vasotropic;  
KW antirheumatic; antiproliferative; cytostatic; cardiant; neuroprotective;  
KW cerebroprotective; nootropic; antibacterial; virucide; fungicide; cancer;  
KW ophthalmological; vulnery; gene therapy; autoimmune disease; neoplasm;  
KW hyperproliferative disorder; breast; liver; cardiovascular disorder; ss;  
KW cerebrovascular disorder; nervous system disorder; bacterial infection;  
KW fungal infection; viral infection; ocular disorder; endocrine disorder;  
KW gastrointestinal disorder; renal disorder; respiratory disorder;  
KW wound healing; skin aging; organ transplantation; food preservative;  
KW tissue regeneration; anti-infertility.  
XX OS Homo sapiens.  
XX PN WO20015364-A2.  
XX PD 02-AUG-2001.  
XX PF 17-JAN-2001; 2001WO-US01308.  
XX PR 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 14-AUG-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226688.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.

PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0233484.  
PR 27-SEP-2000; 2000US-0233834.  
PR 27-SEP-2000; 2000US-0233836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.

QY	181	ACTGATGCAATTAATCTGCGATACCCATTTTGTGACGAAACCTCTTTCACATGCTTCAGGTA	240
Db	182	ACTGATGCAATTAATCTGCGATACCCATTTTGTGACGAAATTTCTTTCACATGCTTCAGGTA	241
QY	241	GAATACACAGGTACCATTAACCTCTTGGGTAGTGTATTTCTGCCATTAAACAGTGCT	300
Db	242	GAATACACAGGTACCATTAACCTCTTGGGTAGTGTATTTCTGCCATTAAACAGTGCT	301
QY	301	TTGAACCCAAATCTCTATACTCTGACACACAGACCACTTTAAAGAAATGATTCATCGGTTT	360
Db	302	TTGAACCCAAATCTCTATACTCTGACACACAGACCACTTTAAAGAAATGATTCATCGGTTT	361
QY	361	TGCGATACTACTAGACACAAAGAAATCTATGACACGAAAGGTATCAGAAAACATATGCTC	420
Db	362	TGCTATACTACTAGACACAAAGAAATCTATGACACGAAAGGTATCAGAAAACATATGCTC	419
QY	421	CATCATTCATCTGGGGGAAATGTGGCCACTGCGAGGAGATGCCACCTGAGTTAA	474
Db	420	CATCATTCATCTGGGTGAAATGTGGCCACTGCGAGGAGATGCCACCTGAGTTAA	473
RESULT 15			
AAS30187			
ID	AAS30187 standard; cDNA; 530 BP.		
AC	XX		
CC	AAS30187;		
CT	XX		
DT	21-NOV-2001 (first entry)		
XX	XX		
DE	DNA encoding renal and cardiovascular-associated protein, Seq ID 33.		
XX	XX		
KW	Human; antiinflammatory; neuroprotective; immunomodulator; vulnery;		
KW	cardiovascular; cytostatic; nephrotropic; antianaemic; nephritis;		
KW	immunosuppressive; kidney disorder; renal failure; hypertension;		
KW	cardiovascular disorder; myocardial infarction; blood disorder; anaemia;		
KW	blood coagulation disorder; electrolyte imbalance disorder; cancer;		
KW	hyponatremia; hyperkalaemia; neoplastic disorder; nephroma;		
KW	autoimmune disease; inflammatory disease; reproductive system disorder;		
KW	endocrine disorder; neural activity; neurological disorder;		
KW	wound healing; respiratory disorder; ss.		
OS	Homo sapiens.		
XX	XX		
PN	WO200155328-A2.		
XX	XX		
PD	02-AUG-2001.		
XX	XX		
PF	17-JAN-2001; 2001WO-US01359.		
XX	XX		
PR	31-JAN-2000; 2000US-0179065.		
PR	04-FEB-2000; 2000US-0180628.		
PR	24-FEB-2000; 2000US-0184664.		
PR	02-MAR-2000; 2000US-0186350.		
PR	16-MAR-2000; 2000US-0189874.		
PR	17-MAR-2000; 2000US-0190076.		
PR	18-APR-2000; 2000US-0198123.		
PR	19-MAY-2000; 2000US-0205515.		
PR	07-JUN-2000; 2000US-0203467.		
PR	28-JUN-2000; 2000US-0214886.		
PR	30-JUN-2000; 2000US-0215135.		
PR	07-JUL-2000; 2000US-0216647.		
PR	07-JUL-2000; 2000US-0216880.		
PR	11-JUL-2000; 2000US-0217487.		
PR	11-JUL-2000; 2000US-0217496.		
PR	14-JUL-2000; 2000US-0218290.		
PR	26-JUL-2000; 2000US-0220963.		
PR	26-JUL-2000; 2000US-0220964.		
PR	14-AUG-2000; 2000US-0224518.		
PR	14-AUG-2000; 2000US-0224519.		
PR	14-AUG-2000; 2000US-0225213.		
PR	14-AUG-2000; 2000US-0225214.		
PR	14-AUG-2000; 2000US-0225216.		

PR	14-AUG-2000;	2000US-0225267.
PR	14-AUG-2000;	2000US-0225268.
PR	14-AUG-2000;	2000US-0225269.
PR	14-AUG-2000;	2000US-0225270.
PR	14-AUG-2000;	2000US-0225271.
PR	14-AUG-2000;	2000US-0225272.
PR	14-AUG-2000;	2000US-0225273.
PR	14-AUG-2000;	2000US-0225274.
PR	14-AUG-2000;	2000US-0225275.
PR	14-AUG-2000;	2000US-0225276.
PR	14-AUG-2000;	2000US-0225277.
PR	14-AUG-2000;	2000US-0225278.
PR	14-AUG-2000;	2000US-0225279.
PR	14-AUG-2000;	2000US-0225280.
PR	22-AUG-2000;	2000US-0226681.
PR	22-AUG-2000;	2000US-0226682.
PR	22-AUG-2000;	2000US-0226683.
PR	22-AUG-2000;	2000US-0227182.
PR	23-AUG-2000;	2000US-0227009.
PR	30-AUG-2000;	2000US-0228924.
PR	01-SEP-2000;	2000US-0229287.
PR	01-SEP-2000;	2000US-0229343.
PR	01-SEP-2000;	2000US-0229344.
PR	01-SEP-2000;	2000US-0229345.
PR	05-SEP-2000;	2000US-0229503.
PR	05-SEP-2000;	2000US-0229509.
PR	06-SEP-2000;	2000US-0230436.
PR	06-SEP-2000;	2000US-0230437.
PR	08-SEP-2000;	2000US-0231242.
PR	08-SEP-2000;	2000US-0231243.
PR	08-SEP-2000;	2000US-0231244.
PR	08-SEP-2000;	2000US-0231413.
PR	08-SEP-2000;	2000US-0231414.
PR	08-SEP-2000;	2000US-0232030.
PR	08-SEP-2000;	2000US-0232081.
PR	12-SEP-2000;	2000US-0231988.
PR	14-SEP-2000;	2000US-0232397.
PR	14-SEP-2000;	2000US-0232398.
PR	14-SEP-2000;	2000US-0232399.
PR	14-SEP-2000;	2000US-0232400.
PR	14-SEP-2000;	2000US-0232401.
PR	14-SEP-2000;	2000US-0233063.
PR	14-SEP-2000;	2000US-0233064.
PR	14-SEP-2000;	2000US-0233065.
PR	21-SEP-2000;	2000US-0234423.
PR	21-SEP-2000;	2000US-0234424.
PR	21-SEP-2000;	2000US-0234425.
PR	25-SEP-2000;	2000US-0234977.
PR	25-SEP-2000;	2000US-0234998.
PR	26-SEP-2000;	2000US-0235484.
PR	27-SEP-2000;	2000US-0235834.
PR	27-SEP-2000;	2000US-0235836.
PR	29-SEP-2000;	2000US-0236327.
PR	29-SEP-2000;	2000US-0236367.
PR	29-SEP-2000;	2000US-0236368.
PR	29-SEP-2000;	2000US-0236369.
PR	29-SEP-2000;	2000US-0236370.
PR	02-OCT-2000;	2000US-0236802.
PR	02-OCT-2000;	2000US-0237037.
PR	02-OCT-2000;	2000US-0237038.
PR	02-OCT-2000;	2000US-0237039.
PR	02-OCT-2000;	2000US-0237040.
PR	13-OCT-2000;	2000US-0239935.
PR	13-OCT-2000;	2000US-0239937.
PR	20-OCT-2000;	2000US-0241809.
PR	20-OCT-2000;	2000US-0241808.
PR	20-OCT-2000;	2000US-0241807.
PR	20-OCT-2000;	2000US-0241826.
PR	01-NOV-2000;	2000US-0244617.
PR	08-NOV-2000;	2000US-0246474.
PR	08-NOV-2000;	2000US-0246475.
PR	08-NOV-2000;	2000US-0246476.
PR	08-NOV-2000;	2000US-0246477.
PR	08-NOV-2000;	2000US-0246478.
PR	08-NOV-2000;	2000US-0246523.
PR	08-NOV-2000;	2000US-0246524.
PR	08-NOV-2000;	2000US-0246525.

08-NOV-2000;	2000US-0246526;
08-NOV-2000;	2000US-0246527;
08-NOV-2000;	2000US-0246538;
08-NOV-2000;	2000US-0246539;
08-NOV-2000;	2000US-0246542;
08-NOV-2000;	2000US-0246543;
08-NOV-2000;	2000US-0246610;
08-NOV-2000;	2000US-0246611;
08-NOV-2000;	2000US-0246613;
17-NOV-2000;	2000US-0249207;
17-NOV-2000;	2000US-0249208;
17-NOV-2000;	2000US-0249209;
17-NOV-2000;	2000US-0249210;
17-NOV-2000;	2000US-0249211;
17-NOV-2000;	2000US-0249212;
17-NOV-2000;	2000US-0249213;
17-NOV-2000;	2000US-0249214;
17-NOV-2000;	2000US-0249215;
17-NOV-2000;	2000US-0249216;
17-NOV-2000;	2000US-0249217;
17-NOV-2000;	2000US-0249218;
17-NOV-2000;	2000US-0249244;
17-NOV-2000;	2000US-0249245;
17-NOV-2000;	2000US-0249284;
17-NOV-2000;	2000US-0249285;
17-NOV-2000;	2000US-0249287;
17-NOV-2000;	2000US-0249299;
17-NOV-2000;	2000US-0249300;
17-NOV-2000;	2000US-0249301;
01-DEC-2000;	2000US-0250160;
01-DEC-2000;	2000US-0250391;
05-DEC-2000;	2000US-0251030;
05-DEC-2000;	2000US-0251198;
05-DEC-2000;	2000US-0256719;
06-DEC-2000;	2000US-0257149;
08-DEC-2000;	2000US-0251185;
08-DEC-2000;	2000US-0251868;
08-DEC-2000;	2000US-0251989;
08-DEC-2000;	2000US-0251990;
11-DEC-2000;	2000US-0254097;
05-JAN-2001;	2001US-0259598;

(HUMA - ) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI: 2001-488787/53.

P-PSDB; AAU18656.

New polynucleotides and polypeptides, useful for diagnosing, treating, preventing or prognosing e.g. kidney, cardiovascular, blood, electrolyte imbalance or neoplastic disorders, autoimmune diseases, cancers -

Claim 1; SEQ ID No 33; 506pp; English.

The invention relates to novel nucleic acids and polypeptides useful for diagnosing, treating, preventing and/or prognosing disorders related to these polypeptides. The polynucleotides are especially useful in the diagnosis, prognosis, prevention and/or treatment of diseases which include kidney disorders (e.g. renal failure or nephritis), cardiovascular disorders (e.g. hypertension or myocardial infarction), blood disorders (e.g. anaemia or blood coagulation disorders), electrolyte imbalance disorders (e.g. hyponatraemia or hyperkalaemia), neoplastic disorders (e.g. nephroma or renal cell cancer), autoimmune diseases, cancers, inflammatory diseases, reproductive system disorders, endocrine disorders, neural activity and neurological disorders, wound healing and respiratory disorders. AAS30165-AAS30251 represent the novel human renal and cardiovascular-associated nucleic acid sequences of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at: [ftp.wipo.int/pub/published\\_pct\\_sequences](http://ftp.wipo.int/pub/published_pct_sequences).

XX 3



Search completed: September 25, 2003, 13:21:51  
Job time : 225 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 25, 2003, 13:58:14 ; Search time 190 Seconds  
(without alignments)  
6210.625 Million cell updates/sec

Title: US-10-049-568-1  
Perfect score: 474  
Sequence: 1 gccagattattcagtgcc.....ggagatgccacctgagittaa 474

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1678620 seqs, 1244745471 residues

Total number of hits satisfying chosen parameters: 3357240

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published\_Applications\_NA:\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09D\_PUBCOMB.seq:\*
- 13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*
- 16: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*
- 17: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	439.2	92.7	1089	14	US-10-073-885-20	Sequence 20, Appl
2	439.2	92.7	1804	10	US-09-895-686-10	Sequence 10, Appl
3	439.2	92.7	2274	14	US-10-225-567A-622	Sequence 622, App
4	438.8	92.6	530	9	US-09-764-853-137	Sequence 137, App
5	438.8	92.6	530	9	US-09-764-853-137	Sequence 986, App
6	438.8	92.6	530	11	US-09-989-442-33	Sequence 33, Appl
7	438.8	92.6	530	12	US-09-764-886-18	Sequence 18, Appl
8	438.8	92.6	530	14	US-10-073-885-25	Sequence 25, Appl
9	438.8	92.6	530	14	US-10-103-313-89	Sequence 89, Appl
10	438.8	92.6	530	14	US-10-073-885-47	Sequence 47, Appl
11	437.6	92.3	1126	14	US-10-313-542-255	Sequence 255, App
12	436.6	92.1	612	10	US-09-895-686-37	Sequence 37, Appl
13	171.8	36.2	1068	12	US-10-321-807-17	Sequence 17, Appl
14	171.8	36.2	2049	11	US-09-928-175-11	Sequence 11, Appl
15	171.8	36.2	2142	11	US-09-965-536A-5	Sequence 5, Appl
16	171.8	36.2	2193	10	US-09-928-175-6	Sequence 6, Appl

1010495568

17	171.8	36.2	2214	11	US-09-965-536A-1	Sequence 1, Appl
18	171.8	36.2	2265	10	US-09-928-175-1	Sequence 1, Appl
19	171.8	36.2	2436	14	US-10-229-735-2	Sequence 2, Appl
20	171.8	36.2	2838	14	US-10-222-668-1	Sequence 1, Appl
21	156.8	33.1	321	11	US-09-930-312-1	Sequence 1, Appl
22	146.2	30.8	2214	10	US-09-928-175-19	Sequence 19, Appl
23	146.2	30.8	2539	14	US-10-229-735-4	Sequence 4, Appl
24	130.4	27.5	719	12	US-09-814-353-18321	Sequence 18321, A
25	113	23.8	677	12	US-09-814-353-5651	Sequence 5651, Ap
26	113	23.8	677	12	US-09-814-353-11937	Sequence 11937, A
27	107.2	22.6	432	10	US-09-895-686-62	Sequence 62, Appl
28	106.8	22.5	1018	11	US-09-782-974C-65	Sequence 65, Appl
29	48.8	10.3	1015	14	US-10-270-333-20	Sequence 20, Appl
30	44	9.3	4292	14	US-10-102-806-179	Sequence 179, App
31	44	9.3	5516	14	US-10-153-688-271	Sequence 271, App
32	39.8	8.4	113515	12	US-10-311-455-2147	Sequence 2147, Ap
33	38.8	8.2	12138	12	US-10-311-455-1602	Sequence 1602, Ap
34	38	8.0	444	13	US-10-027-632-62629	Sequence 62629, A
35	38	8.0	444	13	US-10-027-632-62630	Sequence 62630, A
36	38	8.0	444	13	US-10-027-632-306540	Sequence 306540, A
37	38	8.0	444	13	US-10-027-632-306541	Sequence 306541, A
38	37.8	8.0	1265	9	US-09-810-936-169	Sequence 169, App
39	37.8	8.0	1265	10	US-09-429-755-169	Sequence 169, App
40	37.8	8.0	1265	10	US-09-924-400-169	Sequence 169, App
41	37.8	8.0	1265	14	US-10-212-679-169	Sequence 169, App
42	37.8	8.0	1646	9	US-09-810-936-290	Sequence 290, App
43	37.8	8.0	1646	10	US-09-429-755-290	Sequence 290, App
44	37.8	8.0	1646	10	US-09-924-400-290	Sequence 290, App
45	37.8	8.0	1646	14	US-10-212-679-290	Sequence 290, App

ALIGNMENTS

RESULT 1  
US-10-073-885-20  
; Sequence 20, Application US/10073885  
; Publication No. US20030096346A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PJ203C1  
; CURRENT APPLICATION NUMBER: US/10/073,885  
; CURRENT FILING DATE: 2002-02-14  
; Prior Application removed - See file Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 115  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 20  
; LENGTH: 1089  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-073-885-20

Query Match	92.7%	Score 439.2;	DB 14;	Length 1089;
Best Local Similarity	96.8%	Pred. No. 7.8e-111;		
Matches 459;	Conservative	0;	Mismatches 13;	Indels 2;
Gaps	1;			
QY	1	GCCAGATTATTTCAGTGGCAATTTTCTGTATTATTGTCGCGCATTTATCATCATATA	60	
DB	20	GCCAGATTATTTCAGTGGCAATTTTCTGTATTATTGTCGCGCATTTATCATCATATA	79	
QY	61	GTTCCTTCTATGGAAGCATGTTTATAGTGTTCATCAAAAGTCCCAACTGAA	120	
DB	80	GTTCCTTCTATGGAAGCATGTTTATAGTGTTCATCAAAAGTCCCAACTGAA	139	
QY	121	ATACGATTCAGTATTAAGAGAGATGATCCCTGCCAAGTTCCTTCCTATAGTATT	180	
DB	140	ATACGATTCAGTATTAAGAGAGATGATCCCTGCCAAGTTCCTTCCTATAGTATT	199	
QY	181	ACTGATGATTCATGCTGGATACCCATTTTTFAGCGAAACCTCTTCACTGCTTCAGTA	240	
DB	200	ACTGATGATTCATGCTGGATACCCATTTTTFAGCGAAACCTCTTCACTGCTTCAGTA	259	

QY 241 GAAATACCAAGTACCATAACCTCTGGTAGTGGTATTCTGCGCATACAGTGT 300  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 260 GAAATACCAAGTACCATAACCTCTGGTAGTGGTATTCTGCGCATACAGTGT 319  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 301 TTGAACCAATCTCTATACCTCTGACCAAGACCACTTTAAAGAAATGATTCGTTT 360  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 320 TTGAACCAATCTCTATACCTCTGACCAAGACCACTTTAAAGAAATGATTCGTTT 379  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 361 TGGCATACATACAGCAAGAAATCTATGACAGCAAGATGATCAAGAAATATGCTC 420  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 380 TGGTATACATACAGCAAGAAATCTATGACAGCAAGATGATCAAGAAATATGCTC 437  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 421 CATCATTCATCTGGGGGAATGTGGCACTGCGAGAGATGCGACCTGAGTTAA 474  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 438 CATCATTCATCTGGGGGAATGTGGCACTGCGAGAGATGCGACCTGAGTTAA 491  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||

RESULT 2  
 US-09-895-686-10  
 ; Sequence 10, Application US/09895686  
 ; Patent No. US20020106655A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Bandman, Olga  
 ; APPLICANT: Lal, Preeti  
 ; APPLICANT: Tang, Y. Tom  
 ; APPLICANT: Baughn, Mariah R.  
 ; TITLE OF INVENTION: HUMAN GPCR PROTEINS  
 ; FILE REFERENCE: PC-0044 CIP  
 ; CURRENT APPLICATION NUMBER: US/09/895,686  
 ; CURRENT FILING DATE: 2001-06-28  
 ; NUMBER OF SEQ ID NOS: 74  
 ; SOFTWARE: PERL Program  
 ; SEQ ID NO 10  
 ; LENGTH: 1804  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: misc\_feature  
 ; OTHER INFORMATION: Incyte ID No. US20020106655A1 2488822C51  
 US-09-895-686-10

Query Match 92.7%; Score 439.2; DB 10; Length 1804;  
 Best Local Similarity 96.8%; Pred. No. 1.e-110;  
 Matches 459; Conservative 0; Mismatches 13; Indels 2; Gaps 1;  
 QY 1 GCCAGATTATTCAGTGGCAATTTTCTGGTATTAATTTGGCCGCAATTTATCATCA 60  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 792 GCCAGATTATTCAGTGGCAATTTTCTGGTATTAATTTGGCCGCAATTTATCATCA 851  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 61 GTTTTTCCTATGGAAGCATGTTTATAGTGTTCATCAAGTGCCTAACGACTGAA 120  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 852 GTTTTTCCTATGGAAGCATGTTTATAGTGTTCATCAAGTGCCTAACGACTGAA 911  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 121 ATACGGAATCAAGTATTAAGAGATGATCCTTGCCAAACGTTTCTTTATAGTATT 180  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 912 ATACGGAATCAAGTATTAAGAGATGATCCTTGCCAAACGTTTCTTTATAGTATT 971  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 181 ACTGATGATATGCTGGATACCATTTTCTGAGCAAGACCTTCTCTGCTTCAGGTA 240  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 972 ACTGATGATATGCTGGATACCATTTTCTGAGCAAGACCTTCTCTGCTTCAGGTA 1031  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 241 GAAATACCAAGTACCATAACCTCTGGTAGTGGTATTCTGCGCATACAGTGT 300  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 1032 GAAATACCAAGTACCATAACCTCTGGTAGTGGTATTCTGCGCATACAGTGT 1091  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 301 TTGAACCAATCTCTATACCTCTGACCAAGACCACTTTAAAGAAATGATTCGTTT 360  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 1092 TTGAACCAATCTCTATACCTCTGACCAAGACCACTTTAAAGAAATGATTCGTTT 1151  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 361 TGGCATACATACAGCAAGAAATCTATGACAGCAAGATGATCAAGAAATATGCTC 420  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 1152 TGGTATACATACAGCAAGAAATCTATGACAGCAAGATGATCAAGAAATATGCTC 1209  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 421 CATCATTCATCTGGGGGAATGTGGCACTGCGAGAGATGCGACCTGAGTTAA 474  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 1210 CATCATTCATCTGGGGGAATGTGGCACTGCGAGAGATGCGACCTGAGTTAA 1263  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||

RESULT 3  
 US-10-225-567A-622  
 ; Sequence 622, Application US/10225567A  
 ; Publication No. US20030113798A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: LifeSpan Biosciences  
 ; APPLICANT: Burner, Joseph P.  
 ; APPLICANT: Roush, Christine L.  
 ; TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTOR  
 ; FILE REFERENCE: 1920-4-4  
 ; CURRENT APPLICATION NUMBER: US/10/225,567A  
 ; CURRENT FILING DATE: 2001-12-19  
 ; PRIOR APPLICATION NUMBER: 60/257,144  
 ; PRIOR FILING DATE: 2000-12-19  
 ; NUMBER OF SEQ ID NOS: 2292  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 622  
 ; LENGTH: 2274  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; US-10-225-567A-622

Query Match 92.7%; Score 439.2; DB 14; Length 2274;  
 Best Local Similarity 96.8%; Pred. No. 1.2e-110;  
 Matches 459; Conservative 0; Mismatches 13; Indels 2; Gaps 1;  
 QY 1 GCCAGATTATTCAGTGGCAATTTTCTGGTATTAATTTGGCCGCAATTTATCATCA 60  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 1726 GCCAGATTATTCAGTGGCAATTTTCTGGTATTAATTTGGCCGCAATTTATCATCA 1785  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 61 GTTTTTCCTATGGAAGCATGTTTATAGTGTTCATCAAGTGCCTAACGACTGAA 120  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 1786 GTTTTTCCTATGGAAGCATGTTTATAGTGTTCATCAAGTGCCTAACGACTGAA 1845  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 121 ATACGGAATCAAGTATTAAGAGATGATCCTTGCCAAACGTTTCTTTATAGTATT 180  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 1846 ATACGGAATCAAGTATTAAGAGATGATCCTTGCCAAACGTTTCTTTATAGTATT 1905  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 181 ACTGATGATATGCTGGATACCATTTTCTGAGCAAGACCTTCTCTGCTTCAGGTA 240  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 1906 ACTGATGATATGCTGGATACCATTTTCTGAGCAAGACCTTCTCTGCTTCAGGTA 1965  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 241 GAAATACCAAGTACCATAACCTCTGGTAGTGGTATTCTGCGCATACAGTGT 300  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 1966 GAAATACCAAGTACCATAACCTCTGGTAGTGGTATTCTGCGCATACAGTGT 2025  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 301 TTGAACCAATCTCTATACCTCTGACCAAGACCACTTTAAAGAAATGATTCGTTT 360  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 2026 TTGAACCAATCTCTATACCTCTGACCAAGACCACTTTAAAGAAATGATTCGTTT 2085  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 361 TGGCATACATACAGCAAGAAATCTATGAGCAAGATGATCAAGAAATATGCTC 420  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 2086 TGGTATACATACAGCAAGAAATCTATGAGCAAGATGATCAAGAAATATGCTC 2143  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 421 CATCATTCATCTGGGGGAATGTGGCACTGCGAGAGATGCGACCTGAGTTAA 474  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 2144 CATCATTCATCTGGGGGAATGTGGCACTGCGAGAGATGCGACCTGAGTTAA 2197  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||

RESULT 4  
 US-09-764-853-137  
 ; Sequence 137, Application US/09764853  
 ; Patent No. US20020090672A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Rosen et al.  
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
 ; FILE REFERENCE: P0206

```

; MEDIUM TYPE: Diskette 3.50", 1.44 Mb storage
; COMPUTER: IBM PC/XT/AT
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Word Processor (WordPerfect 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/853,659A
; FILING DATE: Unknown
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: none
; FILING DATE: n/a
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8967 bases
; TYPE: nucleotide
; STRANDEDNESS: double stranded
; TOPOLOGY: linear
US-08-853-659A-9

Query Match
Best Local Similarity 8.1%; Score 38.4; DB 2; Length 8967;
Matches 60; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY 98 AAAGTGCCATAACAGCACTGAATACGGAATCAAGTTAAAAAGAGATGATCCTTGCCA 157
DB 1844 AAATTGCCATAAAAGAGATGAATAAAGATTTAGTATATAAAATCTTGTTCATCGATA 1903

QY 158 AACGTTTTCTTTATAGTATTACTGATCATAT 193
DB 1904 AAGAGTTTTTATGATATTTTACTCTTATACTAT 1939

RESULT 5
US-08-853-659A-64/c
; Sequence 64, Application US/08853659A
; Patent No. 5925522
; GENERAL INFORMATION:
; APPLICANT: Wong, K.K.; Saffer, J.D.
; TITLE OF INVENTION: A Salmonella Sequence, Methods Of Detection
; TITLE OF INVENTION: Of A
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Paul W. Zimmerman
; ADDRESSEE: Intellectual Property Services
; ADDRESSEE: Battelle Memorial Institute
; ADDRESSEE: PNNL P.O. Box 999
; STREET: Washington Way
; CITY: Richland
; STATE: Washington
; COUNTRY: U.S.A.
; ZIP: 99352
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette 3.50", 1.44 Mb storage
; COMPUTER: IBM PC/XT/AT
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Word Processor (WordPerfect 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/853,659A
; FILING DATE: Unknown
; PRIOR APPLICATION DATA:
; CLASSIFICATION: 435
; APPLICATION NUMBER: none
; FILING DATE: n/a
; INFORMATION FOR SEQ ID NO: 67:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8967 bases
; TYPE: nucleotide
; STRANDEDNESS: single stranded
; TOPOLOGY: linear
US-08-853-659A-67

Query Match
Best Local Similarity 8.1%; Score 38.4; DB 2; Length 8967;
Matches 37; Conservative 23; Mismatches 36; Indels 0; Gaps 0;

QY 98 AAAGTGCCATAACAGCACTGAATACGGAATCAAGTTAAAAAGAGATGATCCTTGCCA 157
DB 1844 AAUUGCCAUAAAAGAGAGUAAAAGAUUAGUAAUAAAACUUGUUGUACGAUA 1903

QY 158 AACGTTTTCTTTATAGTATTACTGATCATAT 193
DB 1904 AAGAGUUUUUAGUAAUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUU 1939

RESULT 7
US-08-853-659A-2/c
; Sequence 2, Application US/08853659A
; Patent No. 5925522
; GENERAL INFORMATION:
; APPLICANT: Wong, K.K.; Saffer, J.D.
; TITLE OF INVENTION: A Salmonella Sequence, Methods Of Detection
; TITLE OF INVENTION: Of A

```

```

; TITLE OF INVENTION: Salmonella Sequence, And Methods Of Detection Of Salmonella
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Paul W. Zimmerman
; ADDRESSEE: Intellectual Property Services
; ADDRESSEE: Battelle Memorial Institute
; ADDRESSEE: PNNL P.O. Box 999
; STREET: Washington Way
; CITY: Richland
; STATE: Washington
; COUNTRY: U.S.A.
; ZIP: 99352
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette 3.50", 1.44 Mb storage
; COMPUTER: IBM PC/XT/AT
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Word Processor (WordPerfect 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/853,659A
; FILING DATE: Unknown
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: none
; FILING DATE: n/a
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 24701 bases
; TYPE: nucleotide
; STRANDEDNESS: double stranded
; TOPOLOGY: linear
; US-08-853-659A-2

Query Match      8.1%; Score 38.4; DB 2; Length 24701;
Best Local Similarity 62.5%; Pred. No. 0.68;
Matches 60; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY 98 AAGTGCCTAATACGACACTGAATACGGAATCAAGTTAAAAAGAGAGATCCTTGCCA 157
DB 22858 AATATGCCATAAAAGAGATGAAAAAAGTATTAGTATATAAAAAATCTGTTCATCGATA 22799
QY 158 AAGCTTTTCTTTTATAGTATTACTGATGCATTAT 193
DB 22798 AAGAGTTTTTATGATATTTTACTCTTACTAT 22763

RESULT 8
US-08-853-659A-3
; Sequence 3, Application US/08853659A
; Patent No. 5925522
; GENERAL INFORMATION:
; APPLICANT: Wong, K.K.; Saffer, J.D.
; TITLE OF INVENTION: A Salmonella Sequence, Methods Of Detection
; TITLE OF INVENTION: Of A
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Paul W. Zimmerman
; ADDRESSEE: Intellectual Property Services
; ADDRESSEE: Battelle Memorial Institute
; ADDRESSEE: PNNL P.O. Box 999
; STREET: Washington Way
; CITY: Richland
; STATE: Washington
; COUNTRY: U.S.A.
; ZIP: 99352
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette 3.50", 1.44 Mb storage
; COMPUTER: IBM PC/XT/AT
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Word Processor (WordPerfect 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/853,659A
; FILING DATE: Unknown
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: none
; FILING DATE: n/a
; INFORMATION FOR SEQ ID NO: 60:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 24701 bases
; TYPE: nucleotide
; STRANDEDNESS: single stranded
; TOPOLOGY: linear
; US-08-853-659A-60

Query Match      8.1%; Score 38.4; DB 2; Length 24701;
Best Local Similarity 62.5%; Pred. No. 0.68;
Matches 60; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY 98 AAGTGCCTAATACGACACTGAATACGGAATCAAGTTAAAAAGAGAGATCCTTGCCA 157
DB 22858 AATATGCCATAAAAGAGATGAAAAAAGTATTAGTATATAAAAAATCTGTTCATCGATA 22799
QY 158 AAGCTTTTCTTTTATAGTATTACTGATGCATTAT 193
DB 22798 AAGAGTTTTTATGATATTTTACTCTTACTAT 22763

RESULT 9
US-08-853-659A-60/C
; Sequence 60, Application US/08853659A
; Patent No. 5925522
; GENERAL INFORMATION:
; APPLICANT: Wong, K.K.; Saffer, J.D.
; TITLE OF INVENTION: A Salmonella Sequence, Methods Of Detection
; TITLE OF INVENTION: Of A
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Paul W. Zimmerman
; ADDRESSEE: Intellectual Property Services
; ADDRESSEE: Battelle Memorial Institute
; ADDRESSEE: PNNL P.O. Box 999
; STREET: Washington Way
; CITY: Richland
; STATE: Washington
; COUNTRY: U.S.A.
; ZIP: 99352
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette 3.50", 1.44 Mb storage
; COMPUTER: IBM PC/XT/AT
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Word Processor (WordPerfect 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/853,659A
; FILING DATE: Unknown
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: none
; FILING DATE: n/a
; INFORMATION FOR SEQ ID NO: 60:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 24701 bases
; TYPE: nucleotide
; STRANDEDNESS: single stranded
; TOPOLOGY: linear
; US-08-853-659A-60

Query Match      8.1%; Score 38.4; DB 2; Length 24701;
Best Local Similarity 62.5%; Pred. No. 0.68;
Matches 60; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY 98 AAGTGCCTAATACGACACTGAATACGGAATCAAGTTAAAAAGAGAGATCCTTGCCA 157
DB 22858 AATATGCCATAAAAGAGATGAAAAAAGTATTAGTATATAAAAAATCTGTTCATCGATA 22799
QY 158 AAGCTTTTCTTTTATAGTATTACTGATGCATTAT 193
DB 22798 AAGAGTTTTTATGATATTTTACTCTTACTAT 22763

```

```

; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: none
; FILING DATE: n/a
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 24701 bases
; TYPE: nucleotide
; STRANDEDNESS: double stranded
; TOPOLOGY: linear
; US-08-853-659A-3

Query Match      8.1%; Score 38.4; DB 2; Length 24701;
Best Local Similarity 62.5%; Pred. No. 0.68;
Matches 60; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY 98 AAGTGCCTAATACGACACTGAATACGGAATCAAGTTAAAAAGAGAGATCCTTGCCA 157
DB 1844 AATATGCCATAAAAGAGATGAAAAAAGTATTAGTATATAAAAAATCTGTTCATCGATA 1903
QY 158 AAGCTTTTCTTTTATAGTATTACTGATGCATTAT 193
DB 1904 AAGAGTTTTTATGATATTTTACTCTTACTAT 1939

RESULT 9
US-08-853-659A-60/C
; Sequence 60, Application US/08853659A
; Patent No. 5925522
; GENERAL INFORMATION:
; APPLICANT: Wong, K.K.; Saffer, J.D.
; TITLE OF INVENTION: A Salmonella Sequence, Methods Of Detection
; TITLE OF INVENTION: Of A
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Paul W. Zimmerman
; ADDRESSEE: Intellectual Property Services
; ADDRESSEE: Battelle Memorial Institute
; ADDRESSEE: PNNL P.O. Box 999
; STREET: Washington Way
; CITY: Richland
; STATE: Washington
; COUNTRY: U.S.A.
; ZIP: 99352
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette 3.50", 1.44 Mb storage
; COMPUTER: IBM PC/XT/AT
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Word Processor (WordPerfect 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/853,659A
; FILING DATE: Unknown
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: none
; FILING DATE: n/a
; INFORMATION FOR SEQ ID NO: 60:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 24701 bases
; TYPE: nucleotide
; STRANDEDNESS: single stranded
; TOPOLOGY: linear
; US-08-853-659A-60

Query Match      8.1%; Score 38.4; DB 2; Length 24701;
Best Local Similarity 62.5%; Pred. No. 0.68;
Matches 60; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY 98 AAGTGCCTAATACGACACTGAATACGGAATCAAGTTAAAAAGAGAGATCCTTGCCA 157
DB 22858 AATATGCCATAAAAGAGATGAAAAAAGTATTAGTATATAAAAAATCTGTTCATCGATA 22799
QY 158 AAGCTTTTCTTTTATAGTATTACTGATGCATTAT 193
DB 22798 AAGAGTTTTTATGATATTTTACTCTTACTAT 22763

```

; PRIOR APPLICATION NUMBER: 60/241,808  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/241,826  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/241,786  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/241,221  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/246,475  
; PRIOR FILING DATE: 2000-11-08  
; PRIOR APPLICATION NUMBER: 60/231,243  
; PRIOR FILING DATE: 2000-09-08  
; PRIOR APPLICATION NUMBER: 60/233,065  
; PRIOR FILING DATE: 2000-09-14

Query Match 92.6%; Score 438.8; DB 11; Length 530;  
Best Local Similarity 96.6%; Pred. No. 6.9e-111;  
Matches 458; Conservative 1; Mismatches 13; Indels 2; Gaps 1;

QY 1 GCCCAGATTATTCAGTGGCAATTTTCTTGGTATTAATTTGGCGCAATTTATCATCATA 60  
DB 2 GCCCAGATTATTCAGTGGCAATTTTCTTGGTATTAATTTGGCGCAATTTATCATCATA 61  
QY 61 GTTTTTCCCTATGGAAGCATGTTTATAGTGTTCATCAAAAGTGCCATAACAGCAACTGAA 120  
DB 62 GTTTTTCCCTATGGAAGCATGTTTATAGTGTTCATCAAAAGTGCCATAACAGCAACTGAA 121  
QY 121 ATACGGAATCAAGTTAAAGAGAGATGATCCTTGGCAAGTGGTATTAATTTATAGTATTT 180  
DB 122 ATACGGAATCAAGTTAAAGAGAGATGATCCTTGGCAAGTGGTATTAATTTATAGTATTT 181  
QY 181 ACTGATGCATATGCTGGATACCCATTTTGTAGCGAAACCTCTTCTCACTGCTTCAGGTA 240  
DB 182 ACTGATGCATATGCTGGATACCCATTTTGTAGCGAAACCTCTTCTCACTGCTTCAGGTA 241  
QY 241 GAAATACAGGTACCATACCTCTTGGGTAGTGGTATTTCTGCCATTTAACAGTGTCT 300  
DB 242 GAAATACAGGTACCATACCTCTTGGGTAGTGGTATTTCTGCCATTTAACAGTGTCT 301  
QY 301 TTGAACCAATCTCTATCTGACCAAGACCATTTAAAGAAATGATTCATCGGTTT 360  
DB 302 TTGAACCAATCTCTATCTGACCAAGACCATTTAAAGAAATGATTCATCGGTTT 361  
QY 361 TGGCATTAACACAGCAAAAGAAATCTATGACAGCAAGAGTATCAGAAACATATGCTC 420  
DB 362 TGGTATTAACACAGCAAAAGAAATCTATGACAGCAAGAG--TCAGAAACATATGCTC 419  
QY 421 CATCATTCATCTGGGGGAAATGTGGCCACTGCAGGAGATGCCACCTGAGTTAA 474  
DB 420 CATCATTCATCTGGGGGAAATGTGGCCACTGCAGGAGATGCCACCTGAGTTAA 473

RESULT 7  
US-09-764-886-18  
; Sequence 18, Application US/09764886  
; Publication No. US20030139327A9  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PZ02  
; CURRENT APPLICATION NUMBER: US/09/764,886  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - consult PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 88  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 18  
; LENGTH: 530  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-764-886-18

Query Match 92.6%; Score 438.8; DB 12; Length 530;  
Best Local Similarity 96.6%; Pred. No. 6.9e-111;

Matches 458; Conservative 1; Mismatches 13; Indels 2; Gaps 1;  
QY 1 GCCCAGATTATTCAGTGGCAATTTTCTTGGTATTAATTTGGCGCAATTTATCATCATA 60  
DB 2 GCCCAGATTATTCAGTGGCAATTTTCTTGGTATTAATTTGGCGCAATTTATCATCATA 61  
QY 61 GTTTTTCCCTATGGAAGCATGTTTATAGTGTTCATCAAAAGTGCCATAACAGCAACTGAA 120  
DB 62 GTTTTTCCCTATGGAAGCATGTTTATAGTGTTCATCAAAAGTGCCATAACAGCAACTGAA 121  
QY 121 ATACGGAATCAAGTTAAAGAGAGATGATCCTTGGCAAGTGGTATTAATTTATAGTATTT 180  
DB 122 ATACGGAATCAAGTTAAAGAGAGATGATCCTTGGCAAGTGGTATTAATTTATAGTATTT 181  
QY 181 ACTGATGCATATGCTGGATACCCATTTTGTAGCGAAACCTCTTCTCACTGCTTCAGGTA 240  
DB 182 ACTGATGCATATGCTGGATACCCATTTTGTAGCGAAACCTCTTCTCACTGCTTCAGGTA 241  
QY 241 GAAATACAGGTACCATACCTCTTGGGTAGTGGTATTTCTGCCATTTAACAGTGTCT 300  
DB 242 GAAATACAGGTACCATACCTCTTGGGTAGTGGTATTTCTGCCATTTAACAGTGTCT 301  
QY 301 TTGAACCAATCTCTATCTGACCAAGACCATTTAAAGAAATGATTCATCGGTTT 360  
DB 302 TTGAACCAATCTCTATCTGACCAAGACCATTTAAAGAAATGATTCATCGGTTT 361  
QY 361 TGGCATTAACACAGCAAAAGAAATCTATGACAGCAAGAGTATCAGAAACATATGCTC 420  
DB 362 TGGTATTAACACAGCAAAAGAAATCTATGACAGCAAGAG--TCAGAAACATATGCTC 419  
QY 421 CATCATTCATCTGGGGGAAATGTGGCCACTGCAGGAGATGCCACCTGAGTTAA 474  
DB 420 CATCATTCATCTGGGGGAAATGTGGCCACTGCAGGAGATGCCACCTGAGTTAA 473

RESULT 8  
US-10-073-865-25  
; Sequence 25, Application US/10073865  
; Publication No. US20030044904A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PZ09C1  
; CURRENT APPLICATION NUMBER: US/10/073,865  
; CURRENT FILING DATE: 2002-02-14  
; Prior Application removed - See file Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 154  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 25  
; LENGTH: 530  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-073-865-25  
Query Match 92.6%; Score 438.8; DB 14; Length 530;  
Best Local Similarity 96.6%; Pred. No. 6.9e-111;  
Matches 458; Conservative 1; Mismatches 13; Indels 2; Gaps 1;  
QY 1 GCCCAGATTATTCAGTGGCAATTTTCTTGGTATTAATTTGGCGCAATTTATCATCATA 60  
DB 2 GCCCAGATTATTCAGTGGCAATTTTCTTGGTATTAATTTGGCGCAATTTATCATCATA 61  
QY 61 GTTTTTCCCTATGGAAGCATGTTTATAGTGTTCATCAAAAGTGCCATAACAGCAACTGAA 120  
DB 62 GTTTTTCCCTATGGAAGCATGTTTATAGTGTTCATCAAAAGTGCCATAACAGCAACTGAA 121  
QY 121 ATACGGAATCAAGTTAAAGAGAGATGATCCTTGGCAAGTGGTATTAATTTATAGTATTT 180  
DB 122 ATACGGAATCAAGTTAAAGAGAGATGATCCTTGGCAAGTGGTATTAATTTATAGTATTT 181  
QY 181 ACTGATGCATATGCTGGATACCCATTTTGTAGCGAAACCTCTTCTCACTGCTTCAGGTA 240  
DB 182 ACTGATGCATATGCTGGATACCCATTTTGTAGCGAAACCTCTTCTCACTGCTTCAGGTA 241

QY 241 GAAATACCCAGGTACATACCTCTTGGGTAGTATGTTGGTATATCTGCCATTACAGTGCT 300  
 Db |||||  
 QY 242 GAAATACCCAGGTACATACCTCTTGGGTAGTATGTTGGTATATCTGCCATTACAGTGCT 301  
 Db |||||  
 QY 301 TTGAACCAATCTCTATCTACTGACACACAGACCAATTTAAAGAAATGATTCATCGGTT 360  
 Db |||||  
 QY 302 TTGAACCAATCTCTATCTACTGACACACAGACCAATTTAAAGAAATGATTCATCGGTT 361  
 Db |||||  
 QY 361 TGGCAATACCTACAGCAAAAGAAATCTATGACAGCAAAAGGTATCAGAAACATATGCTC 420  
 Db |||||  
 QY 362 TGGTATATCTACAGCAAAAGAAATCTATGACAGCAAAAGG--TCAGAAACATATGCTC 419  
 Db |||||  
 QY 421 CATCATCTCTGGGGGAATGTGGCCACTGACAGAGATGCCACCTGAGTTAA 474  
 Db |||||  
 QY 420 CATCATCTCTGGGGGAATGTGGCCACTGACAGAGATGCCACCTGAGTTAA 473  
 Db |||||

RESULT 9  
 US-10-103-313-89  
 ; Sequence 89, Application US/10103313  
 ; Publication No. US20030082758A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Rosen et al.  
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
 ; FILE REFERENCE: P207C1  
 ; CURRENT APPLICATION NUMBER: US/10/103,313  
 ; CURRENT FILING DATE: 2002-03-12  
 ; NUMBER OF SEQ ID NOS: 653  
 ; Prior Application removed - See File Wrapper or Palm  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 89  
 ; LENGTH: 530  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-10-103-313-89

Query Match 92.6%; Score 438.8; DB 14; Length 530;  
 Best Local Similarity 96.6%; Pred. No. 6.9e-111;  
 Matches 458; Conservative 1; Mismatches 13; Indels 2; Gaps 1;  
 QY 1 GCCAGATTATTCAGTGGCAATTTCTTGGTATTAATTTGGCGCATTTATCATCATCA 60  
 Db |||||  
 QY 2 GCCAGATTATTCAGTGGCAATTTCTTGGTATTAATTTGGCGCATTTATCATCATCA 61  
 Db |||||  
 QY 61 GTTTTCCATGGAAGAGATGTTATAGTTCATCAAAAGTGCATACAGCAACTGAA 120  
 Db |||||  
 QY 62 GTTTTCCATGGAAGAGATGTTATAGTTCATCAAAAGTGCATACAGCAACTGAA 121  
 Db |||||  
 QY 121 ATACGGAATCAAGTTAAAGAGAGATGATCCTTGCCTGCAAAAGTCTTTTCTTATAGTAT 180  
 Db |||||  
 QY 122 ATACGGAATCAAGTTAAAGAGAGATGATCCTTGCCTGCAAAAGTCTTTTCTTATAGTAT 181  
 Db |||||  
 QY 181 ACTGATGATATGCTGGATACCACTCTTGGGTAGTATGTTGGTATTAATTTGGCGCATTTATCATCATCA 240  
 Db |||||  
 QY 182 ACTGATGATATGCTGGATACCACTCTTGGGTAGTATGTTGGTATTAATTTGGCGCATTTATCATCATCA 241  
 Db |||||  
 QY 241 GAAATACCCAGGTACATACCTCTTGGGTAGTATGTTGGTATTAATTTGGCGCATTTATCATCATCA 300  
 Db |||||  
 QY 301 TTGAACCAATCTCTATCTACTGACACACAGACCAATTTAAAGAAATGATTCATCGGTT 360  
 Db |||||  
 QY 302 TTGAACCAATCTCTATCTACTGACACACAGACCAATTTAAAGAAATGATTCATCGGTT 361  
 Db |||||  
 QY 361 TGGCAATACCTACAGCAAAAGAAATCTATGACAGCAAAAGGTATCAGAAACATATGCTC 420  
 Db |||||  
 QY 362 TGGTATATCTACAGCAAAAGAAATCTATGACAGCAAAAGG--TCAGAAACATATGCTC 419  
 Db |||||  
 QY 421 CATCATCTCTGGGGGAATGTGGCCACTGACAGAGATGCCACCTGAGTTAA 474  
 Db |||||  
 QY 420 CATCATCTCTGGGGGAATGTGGCCACTGACAGAGATGCCACCTGAGTTAA 473  
 Db |||||

RESULT 10  
 US-10-073-885-47  
 ; Sequence 47, Application US/10073885  
 ; Publication No. US20030096346A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Rosen et al.  
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
 ; FILE REFERENCE: P203C1  
 ; CURRENT APPLICATION NUMBER: US/10/073,885  
 ; CURRENT FILING DATE: 2002-02-14  
 ; Prior Application removed - See file Wrapper or Palm  
 ; NUMBER OF SEQ ID NOS: 116  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 47  
 ; LENGTH: 530  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-10-073-885-47

Query Match 92.6%; Score 438.8; DB 14; Length 530;  
 Best Local Similarity 96.6%; Pred. No. 6.9e-111;  
 Matches 458; Conservative 1; Mismatches 13; Indels 2; Gaps 1;  
 QY 1 GCCAGATTATTCAGTGGCAATTTCTTGGTATTAATTTGGCGCATTTATCATCATCA 60  
 Db |||||  
 QY 2 GCCAGATTATTCAGTGGCAATTTCTTGGTATTAATTTGGCGCATTTATCATCATCA 61  
 Db |||||  
 QY 61 GTTTTCCATGGAAGAGATGTTATAGTTCATCAAAAGTGCATACAGCAACTGAA 120  
 Db |||||  
 QY 62 GTTTTCCATGGAAGAGATGTTATAGTTCATCAAAAGTGCATACAGCAACTGAA 121  
 Db |||||  
 QY 121 ATACGGAATCAAGTTAAAGAGAGATGATCCTTGCCTGCAAAAGTCTTTTCTTATAGTAT 180  
 Db |||||  
 QY 122 ATACGGAATCAAGTTAAAGAGAGATGATCCTTGCCTGCAAAAGTCTTTTCTTATAGTAT 181  
 Db |||||  
 QY 181 ACTGATGATATGCTGGATACCACTCTTGGGTAGTATGTTGGTATTAATTTGGCGCATTTATCATCATCA 240  
 Db |||||  
 QY 182 ACTGATGATATGCTGGATACCACTCTTGGGTAGTATGTTGGTATTAATTTGGCGCATTTATCATCATCA 241  
 Db |||||  
 QY 241 GAAATACCCAGGTACATACCTCTTGGGTAGTATGTTGGTATTAATTTGGCGCATTTATCATCATCA 300  
 Db |||||  
 QY 242 GAAATACCCAGGTACATACCTCTTGGGTAGTATGTTGGTATTAATTTGGCGCATTTATCATCATCA 301  
 Db |||||  
 QY 301 TTGAACCAATCTCTATCTACTGACACACAGACCAATTTAAAGAAATGATTCATCGGTT 360  
 Db |||||  
 QY 302 TTGAACCAATCTCTATCTACTGACACACAGACCAATTTAAAGAAATGATTCATCGGTT 361  
 Db |||||  
 QY 361 TGGCAATACCTACAGCAAAAGAAATCTATGACAGCAAAAGGTATCAGAAACATATGCTC 420  
 Db |||||  
 QY 362 TGGTATATCTACAGCAAAAGAAATCTATGACAGCAAAAGG--TCAGAAACATATGCTC 419  
 Db |||||  
 QY 421 CATCATCTCTGGGGGAATGTGGCCACTGACAGAGATGCCACCTGAGTTAA 474  
 Db |||||  
 QY 420 CATCATCTCTGGGGGAATGTGGCCACTGACAGAGATGCCACCTGAGTTAA 473  
 Db |||||

RESULT 11  
 US-10-313-542-255  
 ; Sequence 255, Application US/10313542  
 ; Publication No. US20030120057A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Roopa, Reddy  
 ; APPLICANT: Guegler, Karl, J.  
 ; APPLICANT: Au-Young, Janice  
 ; TITLE OF INVENTION: COMPOSITION FOR DETECTION OF GENES ENCODING MEMBRANE-ASSOCIATE  
 ; FILE REFERENCE: PA-0013 US  
 ; CURRENT APPLICATION NUMBER: US/10/313,542  
 ; CURRENT FILING DATE: 2002-12-05  
 ; PRIOR APPLICATION NUMBER: US/09/495,050  
 ; PRIOR FILING DATE: 2000-01-31  
 ; PRIOR APPLICATION NUMBER: 60/118,318  
 ; PRIOR FILING DATE: 1999-02-01

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 25, 2003, 13:00:53 ; Search time 55 Seconds  
(without alignments)  
3803.920 Million cell updates/sec

Title: US-10-049-568-1

Perfect score: 474

Sequence: 1 gccagatttgcagtgcc.....ggagatgcacctgagtaa 474

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents\_NA.\*
- 1: /cgn2.6/prodata/2/ina/5A\_COMB.seq.\*
  - 2: /cgn2.6/prodata/2/ina/5B\_COMB.seq.\*
  - 3: /cgn2.6/prodata/2/ina/6A\_COMB.seq.\*
  - 4: /cgn2.6/prodata/2/ina/6B\_COMB.seq.\*
  - 5: /cgn2.6/prodata/2/ina/ACTUS\_COMB.seq.\*
  - 6: /cgn2.6/prodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Query Match %	Score	Length	ID	Description
1	437.6	92.3	1126	4	US-09-495-050A-255 Sequence 255, Appl
C 2	38.4	8.1	1797	2	US-08-853-859A-28 Sequence 28, Appl
C 3	38.4	8.1	8967	2	US-08-853-859A-6 Sequence 6, Appl
4	38.4	8.1	8967	2	US-08-853-859A-9 Sequence 9, Appl
C 5	38.4	8.1	8967	2	US-08-853-859A-64 Sequence 64, Appl
C 6	38.4	8.1	8967	2	US-08-853-859A-67 Sequence 67, Appl
C 7	38.4	8.1	24701	2	US-08-853-859A-2 Sequence 2, Appl
C 8	38.4	8.1	24701	2	US-08-853-859A-3 Sequence 3, Appl
C 9	38.4	8.1	24701	2	US-08-853-859A-60 Sequence 60, Appl
C 10	38.4	8.1	24701	2	US-08-853-859A-61 Sequence 61, Appl
C 11	37.8	8.0	1265	3	US-09-062-451-169 Sequence 169, Appl
C 12	37.8	8.0	1265	3	US-09-062-451-169 Sequence 169, Appl
C 13	37.8	8.0	1265	4	US-09-598-326-169 Sequence 169, Appl
C 14	37.8	8.0	1265	4	US-09-289-198-169 Sequence 169, Appl
C 15	37.8	8.0	1557	4	US-09-134-001C-1232 Sequence 1232, Ap
C 16	37.8	8.0	1646	3	US-08-991-789A-290 Sequence 290, Appl
C 17	37.8	8.0	1646	4	US-09-062-451-290 Sequence 290, Appl
C 18	37.8	8.0	1646	4	US-09-289-198-290 Sequence 290, Appl
C 19	37.2	7.8	1440	4	US-09-205-258-81 Sequence 81, Appl
C 20	36	7.5	2166	3	US-09-601-198-95 Sequence 95, Appl
C 21	35.4	7.5	836	3	US-09-352-990-7 Sequence 7, Appl
C 22	34.2	7.2	87350	3	US-08-781-891-79 Sequence 79, Appl
C 23	34.2	7.2	87350	3	US-08-781-891-79 Sequence 79, Appl
C 24	34.2	7.2	87350	4	US-09-618-166-79 Sequence 79, Appl
C 25	34.2	7.2	87350	4	US-09-618-166-79 Sequence 79, Appl
C 26	34.2	7.2	87543	4	US-09-791-211-3 Sequence 3, Appl
C 27	34.2	7.2	87543	4	US-09-791-211-3 Sequence 3, Appl

C 28	34.2	7.2	1830121	4	US-09-557-884-1 Sequence 1, Appl
C 29	34.2	7.2	1830121	4	US-09-643-990A-1 Sequence 1, Appl
C 30	33.6	7.1	72804	4	US-09-288-992-7 Sequence 7, Appl
31	33.6	7.1	72804	4	US-09-657-474-7 Sequence 7, Appl
C 32	33.2	7.0	1830121	4	US-09-557-884-1 Sequence 1, Appl
C 33	33.2	7.0	1830121	4	US-09-643-990A-1 Sequence 1, Appl
C 34	33	7.0	615	3	US-08-998-416-186 Sequence 186, App
C 35	33	7.0	636	3	US-08-998-416-1137 Sequence 1137, App
C 36	33	7.0	837	3	US-08-998-416-288 Sequence 288, App
C 37	33	7.0	15462	3	US-08-073-492-1 Sequence 1, Appl
C 38	32.6	6.9	498	4	US-08-936-165A-53 Sequence 53, Appl
C 39	32.2	6.8	8700	2	US-08-332-625-16 Sequence 16, Appl
40	32.2	6.8	8700	2	US-08-456-961A-16 Sequence 16, Appl
41	32.2	6.8	8700	2	US-08-645-193B-16 Sequence 18, Appl
42	31.8	6.7	665	2	US-08-883-795A-36 Sequence 36, Appl
43	31.8	6.7	1300	4	US-08-936-165A-253 Sequence 253, App
44	31.8	6.7	3001	4	US-09-539-333D-215 Sequence 215, App
45	31.8	6.7	3677	2	US-08-390-888A-2 Sequence 2, Appl

#### ALIGNMENTS

#### RESULT 1

US-09-495-050A-255

; Sequence 255, Application US/09495050A

; Patent No. 6492505

; GENERAL INFORMATION:

; APPLICANT: Roopa, Reddy

; APPLICANT: Guegler, Karl, J.

; APPLICANT: Au-Young, Janice

; TITLE OF INVENTION: COMPOSITION FOR DETECTION OF GENES ENCODING MEMBRANE-ASSOCIAT

; FILE REFERENCE: PA-0013 US

; CURRENT APPLICATION NUMBER: US/09/495,050A

; CURRENT FILING DATE: 2000-01-31

; PRIOR APPLICATION NUMBER: 60/118,318

; PRIOR FILING DATE: February 1, 1999

; NUMBER OF SEQ ID NOS: 305

; SOFTWARE: PERL Program

; SEQ ID NO 255

; LENGTH: 1126

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc feature

; OTHER INFORMATION: Incyte ID No. 6492505 3556218CT1

US-09-495-050A-255

Query Match 92.3%; Score 437.6; DB 4; Length 1126;

Best Local Similarity 96.6%; Pred. No. 3.5e-106;

Matches 458; Conservative 0; Mismatches 14; Indels 2; Gaps 1;

QY	1	GCCAGATTTATTCAGTGGCAATTTTCTGGTATTAAATTTGGCGCATTTATCATCATCA 60
Db	114	GCCAGATTTATTCAGTGGCAATTTTCTGGTATTAAATTTGGCGCATTTATCATCA 173
QY	61	GTCTTTTCCTATCGAAGCATGTTTATAGTGTTCATCAAAAGTCCATACACAACTGAA 120
Db	174	GTCTTTTCCTATCGAAGCATGTTTATAGTGTTCATCAAAAGTCCATACACAACTGAA 233
QY	121	ATACGGAATCAAGTAAAGAGATGATCCCTTGGCAAAAGCTTTTCTTTATAGATTTT 180
Db	234	ATACGGAATCAAGTAAAGAGATGATCCCTTGGCAAAAGCTTTTCTTTATAGATTTT 293
QY	181	ACTGATCATATGCTGGATACCCATTTTGTAGCAAAACCTTTTCACTGCTTCAGGTA 240
Db	294	ACTGATCATATGCTGGATACCCATTTTGTAGTGAATTTCTTCTACTGCTTCAGGTA 353
QY	241	GAATATACAGGTACCATACCTCTTGGGTAGTGGTATTCTTCCATTAACAGTGTCT 300
Db	354	GAATATACAGGTACCATACCTCTTGGGTAGTGGTATTCTTCCATTAACAGTGTCT 413
QY	301	TTGAACCAATTTCTTATCTACTCTGACCAAGACCATTTAAAGAAATGATTCATCGGTTT 360



```

Db 414 TGGACCAATCTCTATATCTGACCAAGACCATTTAAGAAATGATTCGCGTTT 473
Qy 361 TGGCACTACTACAGACAAAGAAATCTATGGACAGCAAGGTATCAGAAAACATATGCTC 420
Db 474 TGGTATCTACTACAGACAAAGAAATCTATGGACAGCAAGG--TCAGAAACATATGCTC 531
Qy 421 CATCATTCATCTGGGGGAAATGTGCCCACTGCAGGAGATGCCACCTGAGTTAA 474
Db 532 CATCATTCATCTGGGTGGAATGTGCCCACTGCAGGAGATGCCACCTGAGTTAA 585

RESULT 2
US-08-853-659A-28/c
; Sequence 28, Application US/08853659A
; Patent No. 5925522
; GENERAL INFORMATION:
; APPLICANT: Wong, K.K.; Saffer, J.D.
; TITLE OF INVENTION: A Salmonella Sequence, Methods Of Detection
; TITLE OF INVENTION: Of A
; TITLE OF INVENTION: Salmonella Sequence, And Methods Of Detection
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Paul W. Zimmerman
; ADDRESSEE: Intellectual Property Services
; ADDRESSEE: Battelle Memorial Institute
; ADDRESSEE: PNNL P.O. Box 999
; STREET: Washington Way
; CITY: Richland
; STATE: Washington
; COUNTRY: U.S.A.
; ZIP: 99352
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette 3.50", 1.44 Mb storage
; COMPUTER: IBM PC/XT/AT
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Word Processor (WordPerfect 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/853,659A
; FILING DATE: n/a
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: none
; FILING DATE: n/a
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1797 bases
; TYPE: nucleotide
; STRANDEDNESS: double stranded
; TOPOLOGY: linear
; FEATURE:
; OTHER INFORMATION: the coding nucleotides of SEQ ID
; OTHER INFORMATION: NO:28 correspond to nucleotides 22630 through
; OTHER INFORMATION: 24426 of SEQ ID NO:2
US-08-853-659A-28
Query Match 8.1%; Score 38.4; DB 2; Length 1797;
Best Local Similarity 62.5%; Pred. No. 0.36;
Matches 60; Conservative 0; Mismatches 36; Indels 0; Gaps 0;
Qy 98 AAGTGGCCATAACAGCACTGAATACGGAATCAAGTTAAAGAGAGATGCTTCGCA 157
Db 229 AAATGGCCATAAAGAGATGAAGAAAGATATTAGTATATAAAATCTTGTCATCGATA 170
Qy 158 AACGTTTTCTTTATAGTATTACTGATGCAATAT 193
Db 169 AAGAGTTTTTATGATATTTTTTACTCTTATACTAT 134

RESULT 3
US-08-853-659A-6/c
; Sequence 6, Application US/08853659A
; Patent No. 5925522
; GENERAL INFORMATION:
; APPLICANT: Wong, K.K.; Saffer, J.D.
; TITLE OF INVENTION: A Salmonella Sequence, Methods Of Detection
; TITLE OF INVENTION: Of A
; TITLE OF INVENTION: Salmonella Sequence, And Methods Of Detection
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Paul W. Zimmerman
; ADDRESSEE: Intellectual Property Services
; ADDRESSEE: Battelle Memorial Institute
; ADDRESSEE: PNNL P.O. Box 999
; STREET: Washington Way
; CITY: Richland
; STATE: Washington
; COUNTRY: U.S.A.
; ZIP: 99352
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette 3.50", 1.44 Mb storage
; COMPUTER: IBM PC/XT/AT
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Word Processor (WordPerfect 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/853,659A
; FILING DATE: n/a
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: none
; FILING DATE: n/a
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1797 bases
; TYPE: nucleotide
; STRANDEDNESS: double stranded
; TOPOLOGY: linear
; FEATURE:
; OTHER INFORMATION: the coding nucleotides of SEQ ID
; OTHER INFORMATION: NO:28 correspond to nucleotides 22630 through
; OTHER INFORMATION: 24426 of SEQ ID NO:2
US-08-853-659A-28
Query Match 8.1%; Score 38.4; DB 2; Length 1797;
Best Local Similarity 62.5%; Pred. No. 0.36;
Matches 60; Conservative 0; Mismatches 36; Indels 0; Gaps 0;
Qy 98 AAGTGGCCATAACAGCACTGAATACGGAATCAAGTTAAAGAGAGATGCTTCGCA 157
Db 229 AAATGGCCATAAAGAGATGAAGAAAGATATTAGTATATAAAATCTTGTCATCGATA 170
Qy 158 AACGTTTTCTTTATAGTATTACTGATGCAATAT 193
Db 169 AAGAGTTTTTATGATATTTTTTACTCTTATACTAT 134

```

```

; GENERAL INFORMATION:
; APPLICANT: Wong, K.K.; Saffer, J.D.
; TITLE OF INVENTION: A Salmonella Sequence, Methods Of Detection
; TITLE OF INVENTION: Of A
; TITLE OF INVENTION: Salmonella Sequence, And Methods Of Detection
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Paul W. Zimmerman
; ADDRESSEE: Intellectual Property Services
; ADDRESSEE: Battelle Memorial Institute
; ADDRESSEE: PNNL P.O. Box 999
; STREET: Washington Way
; CITY: Richland
; STATE: Washington
; COUNTRY: U.S.A.
; ZIP: 99352
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette 3.50", 1.44 Mb storage
; COMPUTER: IBM PC/XT/AT
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Word Processor (WordPerfect 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/853,659A
; FILING DATE: n/a
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: none
; FILING DATE: n/a
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8967 bases
; TYPE: nucleotide
; STRANDEDNESS: double stranded
; TOPOLOGY: linear
; FEATURE:
; OTHER INFORMATION: SEQ ID NO:6 corresponds to
; OTHER INFORMATION: nucleotides 15735 through 24701 of SEQ ID NO:2
US-08-853-659A-6
Query Match 8.1%; Score 38.4; DB 2; Length 8967;
Best Local Similarity 62.5%; Pred. No. 0.53;
Matches 60; Conservative 0; Mismatches 36; Indels 0; Gaps 0;
Qy 98 AAGTGGCCATAACAGCACTGAATACGGAATCAAGTTAAAGAGAGATGCTTCGCA 157
Db 7124 AAATGGCCATAAAGAGATGAAGAAAGATATTAGTATATAAAATCTTGTCATCGATA 7065
Qy 158 AACGTTTTCTTTATAGTATTACTGATGCAATAT 193
Db 7064 AAGAGTTTTTATGATATTTTTTACTCTTATACTAT 7029

RESULT 4
US-08-853-659A-9
; Sequence 9, Application US/08853659A
; Patent No. 5925522
; GENERAL INFORMATION:
; APPLICANT: Wong, K.K.; Saffer, J.D.
; TITLE OF INVENTION: A Salmonella Sequence, Methods Of Detection
; TITLE OF INVENTION: Of A
; TITLE OF INVENTION: Salmonella Sequence, And Methods Of Detection
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Paul W. Zimmerman
; ADDRESSEE: Intellectual Property Services
; ADDRESSEE: Battelle Memorial Institute
; ADDRESSEE: PNNL P.O. Box 999
; STREET: Washington Way
; CITY: Richland
; STATE: Washington
; COUNTRY: U.S.A.
; ZIP: 99352
; COMPUTER READABLE FORM:

```

; CURRENT APPLICATION NUMBER: US/09/764,853  
 ; CURRENT FILING DATE: 2001-01-17  
 ; Prior application data removed - consult PALM or file wrapper  
 ; NUMBER OF SEQ ID NOS: 939  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 137  
 ; LENGTH: 530  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-09-764-853-137

Query Match 92.6%; Score 438.8; DB 9; Length 530;  
 Best Local Similarity 96.6%; Pred. No. 6.9e-111;  
 Matches 458; Conservative 1; Mismatches 13; Indels 2; Gaps 1;  
 QY 1 GCCAGATTATTCAGTGGCAATTTCTTGGTATTAATTTGGCGCATTTATCATCATA 60  
 DB 2 GCCAGATTATTCAGTGGCAATTTCTTGGTATTAATTTGGCGCATTTATCATCATA 61  
 QY 61 GTTTTTCCTATGGAAGCATGTTTATAGTGTTCATCAAGTGCCATACAGCAACTGAA 120  
 DB 62 GTTTTTCCTATGGAAGCATGTTTATAGTGTTCATCAAGTGCCATACAGCAACTGAA 121  
 QY 121 ATACGGAATCAAGTTAAAAAGAGATGATCCTTGCCAAACGTTTTTCTTTATAGTATTT 180  
 DB 122 ATACGGAATCAAGTTAAAAAGAGATGATCCTTGCCAAACGTTTTTCTTTATAGTATTT 181  
 QY 181 ACTGATGCATATGCTGGATACCCATTTTGTAGCGAAGACCTCTTCTACTGCTTCAGGTA 240  
 DB 182 ACTGATGCATATGCTGGATACCCATTTTGTAGCGAAGACCTCTTCTACTGCTTCAGGTA 241  
 QY 241 GAATACCCAGTACCATACCTCTTGGGTAGTGTGTTTCTGCTCCATTAACAGTGT 300  
 DB 242 GAATACCCAGTACCATACCTCTTGGGTAGTGTGTTTCTGCTCCATTAACAGTGT 301  
 QY 301 TTGAACCCCAATTTCTCTATCTCTGACCAACAGACCAATTTAAAGAAATGATTCATCGGTTT 360  
 DB 302 TTGAACCCCAATTTCTCTATCTCTGACCAACAGACCAATTTAAAGAAATGATTCATCGGTTT 361  
 QY 361 TGGCATTAACACAGCAAGAAATCTATGACAGCAAGAGTATCAGAAACATATGCTC 420  
 DB 362 TGGTATACTACAGCAAGAAATCTATGACAGCAAGAGTATCAGAAACATATGCTC 419  
 QY 421 CATCATTCATCTGGGGGAAATGTGGCCACTGACAGGAGATGCCACCTGAGTTAA 474  
 DB 420 CATCATTCATCTGGGGGAAATGTGGCCACTGACAGGAGATGCCACCTGAGTTAA 473

RESULT 5  
 US-09-764-877-986  
 ; Sequence 986, Application US/09764877  
 ; Patent No. US20020147140A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Rosen et al.  
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
 ; FILE REFERENCE: PC005  
 ; CURRENT APPLICATION NUMBER: US/09/764,877  
 ; CURRENT FILING DATE: 2001-01-17  
 ; Prior application data removed - refer to PALM or file wrapper  
 ; NUMBER OF SEQ ID NOS: 4031  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 986  
 ; LENGTH: 530  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-09-764-877-986  
 Query Match 92.6%; Score 438.8; DB 10; Length 530;  
 Best Local Similarity 96.6%; Pred. No. 6.9e-111;  
 Matches 458; Conservative 1; Mismatches 13; Indels 2; Gaps 1;  
 QY 1 GCCAGATTATTCAGTGGCAATTTCTTGGTATTAATTTGGCGCATTTATCATCATA 60

DB 2 GCCAGATTATTCAGTGGCAATTTCTTGGTATTAATTTGGCGCATTTATCATCATA 61  
 QY 61 GTTTTTCCTATGGAAGCATGTTTATAGTGTTCATCAAGTGCCATACAGCAACTGAA 120  
 DB 62 GTTTTTCCTATGGAAGCATGTTTATAGTGTTCATCAAGTGCCATACAGCAACTGAA 121  
 QY 121 ATACGGAATCAAGTTAAAAAGAGATGATCCTTGCCAAACGTTTTTCTTTATAGTATTT 180  
 DB 122 ATACGGAATCAAGTTAAAAAGAGATGATCCTTGCCAAACGTTTTTCTTTATAGTATTT 181  
 QY 181 ACTGATGCATATGCTGGATACCCATTTTGTAGCGAAGACCTCTTCTACTGCTTCAGGTA 240  
 DB 182 ACTGATGCATATGCTGGATACCCATTTTGTAGCGAAGACCTCTTCTACTGCTTCAGGTA 241  
 QY 241 GAATACCCAGTACCATACCTCTTGGGTAGTGTGTTTCTGCTCCATTAACAGTGT 300  
 DB 242 GAATACCCAGTACCATACCTCTTGGGTAGTGTGTTTCTGCTCCATTAACAGTGT 301  
 QY 301 TTGAACCCCAATTTCTCTATCTCTGACCAACAGACCAATTTAAAGAAATGATTCATCGGTTT 360  
 DB 302 TTGAACCCCAATTTCTCTATCTCTGACCAACAGACCAATTTAAAGAAATGATTCATCGGTTT 361  
 QY 361 TGGCATTAACACAGCAAGAAATCTATGACAGCAAGAGTATCAGAAACATATGCTC 420  
 DB 362 TGGTATACTACAGCAAGAAATCTATGACAGCAAGAGTATCAGAAACATATGCTC 419  
 QY 421 CATCATTCATCTGGGGGAAATGTGGCCACTGACAGGAGATGCCACCTGAGTTAA 474  
 DB 420 CATCATTCATCTGGGGGAAATGTGGCCACTGACAGGAGATGCCACCTGAGTTAA 473

RESULT 6  
 US-09-989-442-33  
 ; Sequence 33, Application US/09989442  
 ; Publication No. US20030013649A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Rosen et al.  
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
 ; FILE REFERENCE: PJZ08  
 ; CURRENT APPLICATION NUMBER: US/09/989,442  
 ; CURRENT FILING DATE: 2001-11-21  
 ; PRIOR APPLICATION NUMBER: 60/179,065  
 ; PRIOR FILING DATE: 2000-01-31  
 ; PRIOR APPLICATION NUMBER: 60/180,628  
 ; PRIOR FILING DATE: 2000-02-04  
 ; PRIOR APPLICATION NUMBER: 60/214,886  
 ; PRIOR FILING DATE: 2000-06-28  
 ; PRIOR APPLICATION NUMBER: 60/217,487  
 ; PRIOR FILING DATE: 2000-07-11  
 ; PRIOR APPLICATION NUMBER: 60/225,758  
 ; PRIOR FILING DATE: 2000-08-14  
 ; PRIOR APPLICATION NUMBER: 60/220,963  
 ; PRIOR FILING DATE: 2000-07-26  
 ; PRIOR APPLICATION NUMBER: 60/217,496  
 ; PRIOR FILING DATE: 2000-07-11  
 ; PRIOR APPLICATION NUMBER: 60/225,447  
 ; PRIOR FILING DATE: 2000-08-14  
 ; PRIOR APPLICATION NUMBER: 60/218,290  
 ; PRIOR FILING DATE: 2000-07-14  
 ; PRIOR APPLICATION NUMBER: 60/225,757  
 ; PRIOR FILING DATE: 2000-08-14  
 ; PRIOR APPLICATION NUMBER: 60/226,868  
 ; PRIOR FILING DATE: 2000-08-22  
 ; PRIOR APPLICATION NUMBER: 60/216,647  
 ; PRIOR FILING DATE: 2000-07-07  
 ; PRIOR APPLICATION NUMBER: 60/225,267  
 ; PRIOR FILING DATE: 2000-08-14  
 ; PRIOR APPLICATION NUMBER: 60/216,880  
 ; PRIOR FILING DATE: 2000-07-07  
 ; PRIOR APPLICATION NUMBER: 60/225,270  
 ; PRIOR FILING DATE: 2000-08-14  
 ; PRIOR APPLICATION NUMBER: 60/251,869  
 ; PRIOR FILING DATE: 2000-12-08

PRIOR FILING DATE: 2000-11-08	
PRIOR APPLICATION NUMBER: 60/246,532	
PRIOR FILING DATE: 2000-11-08	
PRIOR APPLICATION NUMBER: 60/249,216	
PRIOR FILING DATE: 2000-11-17	
PRIOR APPLICATION NUMBER: 60/249,210	
PRIOR FILING DATE: 2000-11-17	
PRIOR APPLICATION NUMBER: 60/226,681	
PRIOR FILING DATE: 2000-08-22	
PRIOR APPLICATION NUMBER: 60/225,759	
PRIOR FILING DATE: 2000-08-14	
PRIOR APPLICATION NUMBER: 60/225,213	
PRIOR FILING DATE: 2000-08-14	
PRIOR APPLICATION NUMBER: 60/227,182	
PRIOR FILING DATE: 2000-08-22	
PRIOR APPLICATION NUMBER: 60/225,214	
PRIOR FILING DATE: 2000-08-14	
PRIOR APPLICATION NUMBER: 60/235,836	
PRIOR FILING DATE: 2000-09-27	
PRIOR APPLICATION NUMBER: 60/230,438	
PRIOR FILING DATE: 2000-09-06	
PRIOR APPLICATION NUMBER: 60/215,135	
PRIOR FILING DATE: 2000-06-30	
PRIOR APPLICATION NUMBER: 60/225,266	
PRIOR FILING DATE: 2000-08-14	
PRIOR APPLICATION NUMBER: 60/249,218	
PRIOR FILING DATE: 2000-11-17	
PRIOR APPLICATION NUMBER: 60/249,208	
PRIOR FILING DATE: 2000-11-17	
PRIOR APPLICATION NUMBER: 60/249,213	
PRIOR FILING DATE: 2000-11-17	
PRIOR APPLICATION NUMBER: 60/249,212	
PRIOR FILING DATE: 2000-11-17	
PRIOR APPLICATION NUMBER: 60/249,207	
PRIOR FILING DATE: 2000-11-17	
PRIOR APPLICATION NUMBER: 60/249,245	
PRIOR FILING DATE: 2000-11-17	
PRIOR APPLICATION NUMBER: 60/249,244	
PRIOR FILING DATE: 2000-11-17	
PRIOR APPLICATION NUMBER: 60/249,217	
PRIOR FILING DATE: 2000-11-17	
PRIOR APPLICATION NUMBER: 60/249,211	
PRIOR FILING DATE: 2000-11-17	
PRIOR APPLICATION NUMBER: 60/249,215	
PRIOR FILING DATE: 2000-11-17	
PRIOR APPLICATION NUMBER: 60/249,264	
PRIOR FILING DATE: 2000-11-17	
PRIOR APPLICATION NUMBER: 60/249,214	
PRIOR FILING DATE: 2000-11-17	
PRIOR APPLICATION NUMBER: 60/249,297	
PRIOR FILING DATE: 2000-11-17	
PRIOR APPLICATION NUMBER: 60/232,400	
PRIOR FILING DATE: 2000-09-14	
PRIOR APPLICATION NUMBER: 60/231,242	
PRIOR FILING DATE: 2000-09-08	
PRIOR APPLICATION NUMBER: 60/232,081	
PRIOR FILING DATE: 2000-09-08	
PRIOR APPLICATION NUMBER: 60/232,080	
PRIOR FILING DATE: 2000-09-08	
PRIOR APPLICATION NUMBER: 60/231,414	
PRIOR FILING DATE: 2000-09-08	
PRIOR APPLICATION NUMBER: 60/231,244	
PRIOR FILING DATE: 2000-09-08	
PRIOR APPLICATION NUMBER: 60/233,064	
PRIOR FILING DATE: 2000-09-14	
PRIOR APPLICATION NUMBER: 60/233,063	
PRIOR FILING DATE: 2000-09-14	
PRIOR APPLICATION NUMBER: 60/232,397	
PRIOR FILING DATE: 2000-09-14	
PRIOR APPLICATION NUMBER: 60/232,399	
PRIOR FILING DATE: 2000-09-14	
PRIOR APPLICATION NUMBER: 60/232,401	
PRIOR FILING DATE: 2000-09-14	



```
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.419C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 169:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1265 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-062-451-169

Query Match      8.0%; Score 37.8; DB 4; Length 1265;
Best Local Similarity 50.8%; Pred. No. 0.47;
Matches 90; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

QY 240 AGAAATACCAAGGTACCAATACCTCTGGGTAGTGGTATTTCTGCGATTACAGTGC 299
Db 792 AAATAAACAGGCTTCAGATTATCTGGCTTCATATATATTTCTTTTAAAGAAA 733
QY 300 TTGGAACCAATCTCTATCTATCTTGACCAAGACCAATTTAAAGAAATGATTCGTT 359
Db 732 ATATCAACCCATTGTCAATGCATGTTTTCAAAGCAATTAATAGAGGTAACCCCTT 673
QY 360 TTGGCATAACTACAGACAAGAAATCTATGGACAGCAAGGTATCAGAAAACATAT 416
Db 672 TGGAAATTAATACAGAGAAGAAATGATTCACCTTTATGCTATAAAATAATATATAT 616

RESULT 13
US-09-598-326-169/c
; Sequence 169, Application US/09598326
; Patent No. 6423496
; GENERAL INFORMATION:
; APPLICANT: Frudakis, Tony N.
; APPLICANT: Reed, Steven G.
; APPLICANT: Misher, Lynda
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TREATMENT AND DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.419C5
; CURRENT APPLICATION NUMBER: US/09/289,198
; CURRENT FILING DATE: 1999-04-09
; EARLIER APPLICATION NUMBER: US 09/062,451
; EARLIER FILING DATE: 1998-04-17
; EARLIER APPLICATION NUMBER: US 08/991,789
; EARLIER FILING DATE: 1997-12-11
; EARLIER APPLICATION NUMBER: US 08/838,762
; EARLIER FILING DATE: 1997-04-09
; EARLIER APPLICATION NUMBER: PCT/US97/00485
; EARLIER FILING DATE: 1997-01-10
; EARLIER APPLICATION NUMBER: US 08/700,014
; EARLIER FILING DATE: 1996-08-20
; EARLIER APPLICATION NUMBER: US 08/585,392
; EARLIER FILING DATE: 1996-01-01
; NUMBER OF SEQ ID NOS: 312
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 169
; LENGTH: 1265
; TYPE: DNA
; ORGANISM: Homo sapien
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/598,326
FILING DATE: 20-Jun-2000
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Potter, Jane E.R.
REGISTRATION NUMBER: 33,332
REFERENCE/DOCKET NUMBER: 210121.419D1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 169:
SEQUENCE CHARACTERISTICS:
LENGTH: 1265 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 169:
US-09-598-326-169
```

```
Query Match      8.0%; Score 37.8; DB 4; Length 1265;
Best Local Similarity 50.8%; Pred. No. 0.47;
Matches 90; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

QY 240 AGAAATACCAAGGTACCAATACCTCTGGGTAGTGGTATTTCTGCGATTACAGTGC 299
Db 792 AAATAAACAGGCTTCAGATTATCTGGCTTCATATATATTTCTTTTAAAGAAA 733
QY 300 TTGGAACCAATCTCTATCTATCTTGACCAAGACCAATTTAAAGAAATGATTCGTT 359
Db 732 ATATCAACCCATTGTCAATGCATGTTTTCAAAGCAATTAATAGAGGTAACCCCTT 673
QY 360 TTGGCATAACTACAGACAAGAAATCTATGGACAGCAAGGTATCAGAAAACATAT 416
Db 672 TGGAAATTAATACAGAGAAGAAATGATTCACCTTTATGCTATAAAATAATATATAT 616

RESULT 14
US-09-289-198-169/c
; Sequence 169, Application US/09289198
; Patent No. 6586570
; GENERAL INFORMATION:
; APPLICANT: Frudakis, Tony N.
; APPLICANT: Smith, John M.
; APPLICANT: Reed, Steven G.
; APPLICANT: Misher, Lynda
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TREATMENT AND DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.419C5
; CURRENT APPLICATION NUMBER: US/09/289,198
; CURRENT FILING DATE: 1999-04-09
; EARLIER APPLICATION NUMBER: US 09/062,451
; EARLIER FILING DATE: 1998-04-17
; EARLIER APPLICATION NUMBER: US 08/991,789
; EARLIER FILING DATE: 1997-12-11
; EARLIER APPLICATION NUMBER: US 08/838,762
; EARLIER FILING DATE: 1997-04-09
; EARLIER APPLICATION NUMBER: PCT/US97/00485
; EARLIER FILING DATE: 1997-01-10
; EARLIER APPLICATION NUMBER: US 08/700,014
; EARLIER FILING DATE: 1996-08-20
; EARLIER APPLICATION NUMBER: US 08/585,392
; EARLIER FILING DATE: 1996-01-01
; NUMBER OF SEQ ID NOS: 312
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 169
; LENGTH: 1265
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-289-198-169

Query Match      8.0%; Score 37.8; DB 4; Length 1265;
Best Local Similarity 50.8%; Pred. No. 0.47;
Matches 90; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

QY 240 AGAAATACCAAGGTACCAATACCTCTGGGTAGTGGTATTTCTGCGATTACAGTGC 299
Db 792 AAATAAACAGGCTTCAGATTATCTGGCTTCATATATATTTCTTTTAAAGAAA 733
QY 300 TTGGAACCAATCTCTATCTATCTTGACCAAGACCAATTTAAAGAAATGATTCGTT 359
Db 732 ATATCAACCCATTGTCAATGCATGTTTTCAAAGCAATTAATAGAGGTAACCCCTT 673
QY 360 TTGGCATAACTACAGACAAGAAATCTATGGACAGCAAGGTATCAGAAAACATAT 416
Db 672 TGGAAATTAATACAGAGAAGAAATGATTCACCTTTATGCTATAAAATAATATATAT 616

RESULT 15
US-09-134-001C-1232/c
; Sequence 1232, Application US/09134001C
; Patent No. 6380370
```

/ NUMBER OF SEQ ID NOS: 305  
/ SOFTWARE: PERL Program  
/ SEQ ID NO 255  
/ LENGTH: 1126  
/ TYPE: DNA  
/ ORGANISM: Homo sapiens  
/ FEATURE:  
/ NAME/KEY: misc\_feature  
/ OTHER INFORMATION: Incyte ID No. US20030120057A1 3556218CT1  
US-10-313-542-255

Query Match 92.3%; Score 437.6; DB 14; Length 1126;  
Best Local Similarity 96.6%; Pred. No. 2.2e-110;  
Matches 458; Conservative 0; Mismatches 14; Indels 2; Gaps 1;

QY 1 GCCAGATTATTCAGTGGCAATTTTCTGGTATTAAATTTGGCCGGCAATTTATCATCAPA 60  
DB 114 GCCAGATTATTCAGTGGCAATTTTCTGGTATTAAATTTGGCCGGCAATTTATCATCA 173  
QY 61 GTTTTTCCTATGGAAGCATCTTTTATAGTGTTCATCAAGTGCATACAGCAACTGAA 120  
DB 174 GTTTTTCCTATGGAAGCATCTTTTATAGTGTTCATCAAGTGCATACAGCAACTGAA 233  
QY 121 ATACGAATCAAGTTTAAAAAGAGATGATCCTTGGCCAAACGTTTTTCTTTATAGTATT 180  
DB 234 ATACGAATCAAGTTTAAAAAGAGATGATCCTTGGCCAAACGTTTTTCTTTATAGTATT 293  
QY 181 ACTGATGCATTTATGCTGATACCCATTTTGTAGGCAACCTCTTCACTGCTTACAGTA 240  
DB 294 ACTGATGCATTTATGCTGATACCCATTTTGTAGTGAATTTTCTTCACTGCTTACAGTA 353  
QY 241 GAAATACAGGTACCATACCTCTGGGTAGTGTATTTCTGCCATTAAACAGTCT 300  
DB 354 GAAATACAGGTACCATACCTCTGGGTAGTGTATTTCTGCCATTAAACAGTCT 413  
QY 301 TTGAACCAATTTCTATACCTCTGACCAAGACCATTTTAAAGAAATGATTCATCGGTT 360  
DB 414 TTGAACCAATTTCTATACCTCTGACCAAGACCATTTTAAAGAAATGATTCATCGGTT 473  
QY 361 TGCATAACTACAGCAAGAAATCTATGCGACGACCAAGATATCAGAAACATATGCTC 420  
DB 474 TGTATTAATACAGCAAGAAATCTATGCGACGACCAAGATATCAGAAACATATGCTC 531  
QY 421 CATCATTCATCTGGGGGAAATGTGGCCACTGCAGGAGATGCCACCTGAGTTAA 474  
DB 532 CATCATTCATCTGGGTGAATGTGGCCACTGCAGGAGATGCCACCTGAGTTAA 585

## RESULT 12

US-09-895-686-37/c  
/ Sequence 37, Application US/09895686  
/ Patent No. US20020106655A1  
/ GENERAL INFORMATION:  
/ APPLICANT: Bandman, Olga  
/ APPLICANT: Lal, Preeti  
/ APPLICANT: Tang, Y. Tom  
/ APPLICANT: Baughn, Mariah R.  
/ TITLE OF INVENTION: HUMAN GPCR PROTEINS  
/ FILE REFERENCE: PC-0044 CIP  
/ CURRENT APPLICATION NUMBER: US/09/895,686  
/ CURRENT FILING DATE: 2001-06-28  
/ NUMBER OF SEQ ID NOS: 74  
/ SOFTWARE: PERL Program  
/ SEQ ID NO 37  
/ LENGTH: 612  
/ TYPE: DNA  
/ ORGANISM: Homo sapiens  
/ FEATURE:  
/ NAME/KEY: misc\_feature  
/ OTHER INFORMATION: Incyte ID No. US20020106655A1 2488822X308B1  
/ NAME/KEY: unsure  
/ LOCATION: 561  
/ OTHER INFORMATION: a, t, c, g, or other

US-09-895-686-37

Query Match 92.1%; Score 436.6; DB 10; Length 612;  
Best Local Similarity 96.6%; Pred. No. 3e-110;  
Matches 457; Conservative 0; Mismatches 14; Indels 2; Gaps 1;

QY 2 CCAGATTATTCAGTGGCAATTTTCTGGTATTAAATTTGGCCGGCAATTTATCATCATAG 61  
DB 497 CCAGATTATTCAGTGGCAATTTTCTGGTATTAAATTTGGCCGGCAATTTATCATCAG 438  
QY 62 TTTTTCCTATGGAAGCATCTTTTATAGTGTTCATCAAGTGCATACAGCAACTGAAA 121  
DB 437 TTTTTCCTATGGAAGCATCTTTTATAGTGTTCATCAAGTGCATACAGCAACTGAAA 378  
QY 122 TACGGAATCAAGTTTAAAAAGAGATGATCCTTGGCCAAACGTTTTTCTTTATAGTATT 181  
DB 377 TACGGAATCAAGTTTAAAAAGAGATGATCCTTGGCCAAACGTTTTTCTTTATAGTATT 318  
QY 182 CTGATCATTCATCTGGGTGAATGTGGCCACTGCAGGAGATGCCACCTGAGTTAA 241  
DB 317 CTGATCATTCATCTGGGTGAATGTGGCCACTGCAGGAGATGCCACCTGAGTTAA 258  
QY 242 AATACAGGTACCATACCTCTGGGTAGTGTATTTCTGCCATTAAACAGTCTT 301  
DB 257 AATACAGGTACCATACCTCTGGGTAGTGTATTTCTGCCATTAAACAGTCTT 198  
QY 302 TGAACCAATTTCTATACCTCTGACCAAGACCATTTTAAAGAAATGATTCATCGGTTT 361  
DB 197 TGAACCAATTTCTATACCTCTGACCAAGACCATTTTAAAGAAATGATTCATCGGTTT 138  
QY 362 GGCAATACCTACAGCAAGAAATCTATGACAGCAAGATATCAGAAACATATGCTCC 421  
DB 137 GGTATTAATACAGCAAGAAATCTATGACAGCAAGATATCAGAAACATATGCTCC 80  
QY 422 ATCATTCATCTGGGGGAAATGTGGCCACTGCAGGAGATGCCACCTGAGTTAA 474  
DB 79 ATCATTCATCTGGGTGAATGTGGCCACTGCAGGAGATGCCACCTGAGTTAA 27

## RESULT 13

US-10-321-807-17  
/ Sequence 17, Application US/10321807  
/ Publication No. US20030166148A1  
/ GENERAL INFORMATION:  
/ APPLICANT: Chen, Rupong  
/ APPLICANT: Dang, Huong T.  
/ APPLICANT: Lowitz, Kevin P.  
/ TITLE OF INVENTION: No. US20030166148A1-Endogenous, Constitutively Activated Hum  
/ FILE REFERENCE: AREN0086  
/ CURRENT APPLICATION NUMBER: US/10/321,807  
/ CURRENT FILING DATE: 2002-12-16  
/ PRIOR APPLICATION NUMBER: US/09/114,008  
/ PRIOR FILING DATE: 2000-11-16  
/ PRIOR APPLICATION NUMBER: 09/170,496  
/ PRIOR FILING DATE: 1999-11-17  
/ PRIOR APPLICATION NUMBER: PCT/US99/23938  
/ PRIOR FILING DATE: 2000-04-20  
/ PRIOR APPLICATION NUMBER: 60/166,088  
/ PRIOR FILING DATE: 1999-11-17  
/ PRIOR APPLICATION NUMBER: 60/166,099  
/ PRIOR FILING DATE: 1999-11-17  
/ PRIOR APPLICATION NUMBER: 60/166,369  
/ PRIOR FILING DATE: 1999-11-17  
/ PRIOR APPLICATION NUMBER: 60/171,902  
/ PRIOR FILING DATE: 1999-12-23  
/ PRIOR APPLICATION NUMBER: 60/171,901  
/ PRIOR FILING DATE: 1999-12-23  
/ PRIOR APPLICATION NUMBER: 60/171,900  
/ PRIOR FILING DATE: 1999-12-23  
/ PRIOR APPLICATION NUMBER: 60/181,749  
/ PRIOR FILING DATE: 2000-02-11  
/ Remaining Prior Application data removed - See File Wrapper or PALM.

```

; NUMBER OF SEQ ID NOS: 133
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 17
; LENGTH: 1068
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-321-807-17

Query Match          36.2%; Score 171.8; DB 12; Length 1068;
Best Local Similarity 67.4%; Pred. No. 4.1e-37;
Matches 242; Conservative 0; Mismatches 117; Indels 0; Gaps 0;

QY 10 TATTCAGTGGCAATTTTCTGGTATTAATTTGGCCGCAATTTATCATCATAGTTTTC 69
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 568 TATTCCTGTGGAATTTCTAGGCTGGAACCTGCTGCTTTCTCATCATGTTGTTTC 627
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 70 TATGGAACATGTTTATAGTGTTCATCAAAAGTCCATACACAGCACTCAATACGGAAT 129
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 628 TATATTACTATGTCCTGCTCCATTCAAAACCCGCTTGCAGACCAAGAGTAAGGAAT 687
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 130 CAAGTTAAAAAGAGATGATCTGTCGCAAAAGCTTTTCTTTATAGTATTTACTGATGCA 189
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 688 TGTTTTGAAGAGAGTGGCTGTGCAAAATCGTTTCTTTTATAGTGTTCCTGATGCC 747
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 190 TTATGCTGGATACCAATTTTGTAGCGAAACCTCTTCACTGCTTCAGGTAGAAATACCA 249
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 748 ATCTGCTGGATTCCTGTATTTAGTATTAATTCCTTCCCTCTCCGGGTGGAATACCA 807
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 250 GTACCATAACTCTTGGGTAGTATGTTATTCCTCCATTAACAGTGCCTTGAACCCA 309
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 808 GACACAAGTCTCTGGATAGTATTTTCCCTCCAGTTACAGTGCCTTGAATACCA 867
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 310 ATCTCTATATCTGACCAAGACCATTTAAAGAAATGATCAGGTTTGGCATAA 368
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 868 ATCTCTATATCTGACCAACCACTTTTAAAGCAAGTGAACAGTGCCTGACAA 926
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 14
US-09-928-175-11
; Sequence 11, Application US/09928175
; Patent No. US20020123618A1
; GENERAL INFORMATION:
; APPLICANT: Paszty, Christopher J.
; APPLICANT: Gong, Jianhua
; APPLICANT: Daugherty, Betsy
; APPLICANT: Rogers, No. US20020123618Alma
; TITLE OF INVENTION: Leucine-Rich G Protein Coupled Receptor-8 Molecules and
; FILE REFERENCE: Uses Thereof
; CURRENT APPLICATION NUMBER: US/09/928,175
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/224,455
; PRIOR FILING DATE: 2000-08-10
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 2049
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2046)
; NAME/KEY: sig_peptide
; LOCATION: (1)..(108)
US-09-928-175-11

Query Match          36.2%; Score 171.8; DB 10; Length 2049;
Best Local Similarity 67.4%; Pred. No. 5.8e-37;
Matches 242; Conservative 0; Mismatches 117; Indels 0; Gaps 0;

QY 10 TATTCAGTGGCAATTTTCTGGTATTAATTTGGCCGCAATTTATCATCATAGTTTTC 69
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1549 TATTCCTGTGGAATTTCTAGTGTGAACCTGCTGCTTTCTCATCATGTTGTTTC 1608
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

QY 70 TATGGAAGCATGTTTATAGTGTTCATCAAAAGTCCATACAGCACTCAATACGGAAT 129
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1609 TATATTACTATGTCCTGCTCCATTCAAAACCCGCTTGCAGACCAAGAGTAAGGAAT 1668
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 130 CAAGTTAAAAAGAGATGATCTGTCGCAAAAGCTTTTCTTTATAGTATTTACTGATGCA 189
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1659 TGTTTTGAAGAGAGTGGCTGTGCAAAATCGTTTCTTTTATAGTGTTCCTGATGCC 1728
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 190 TTATGCTGGATACCAATTTTGTAGCGAAACCTCTTCACTGCTTCAGGTAGAAATACCA 249
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1729 ATCTGCTGGATTCCTGTATTTAGTATTAATTCCTTCCCTCTCCGGGTGGAATACCA 1788
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 250 GTACCATAACTCTTGGGTAGTATGTTATTCCTCCATTAACAGTGCCTTGAACCCA 309
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1789 GACACAATGACTCTCGGATAGTATTTTCCCTCCAGTTACAGTGCCTTGAATACCA 1848
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 310 ATCTCTATATCTGACCAAGACCATTTAAAGAAATGATTCATCGGTTTGGCATAA 368
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1849 ATCTCTATATCTGACCAACCACTTTTAAAGCAAGTGAACAGCTGCTGACAA 1907
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 15
US-09-965-536A-5
; Sequence 5, Application US/09965536A
; Publication No. US20030027323A1
; GENERAL INFORMATION:
; APPLICANT: FEDER, J. N.
; APPLICANT: MINTIER, G.
; APPLICANT: RAMANATHAN, C. S.
; APPLICANT: HAWKEN, D. R.
; TITLE OF INVENTION: A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR, HGPBMY5,
; FILE REFERENCE: DOO41NP
; CURRENT APPLICATION NUMBER: US/09/965,536A
; CURRENT FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/235,713
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/261,781
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: 60/306,605
; PRIOR FILING DATE: 2001-07-19
; PRIOR APPLICATION NUMBER: 60/310,436
; PRIOR FILING DATE: 2001-08-03
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 2142
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-965-536A-5

Query Match          36.2%; Score 171.8; DB 11; Length 2142;
Best Local Similarity 67.4%; Pred. No. 5.9e-37;
Matches 242; Conservative 0; Mismatches 117; Indels 0; Gaps 0;

QY 10 TATTCAGTGGCAATTTTCTGGTATTAATTTGGCCGCAATTTATCATCATAGTTTTC 69
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1642 TATTCCTGTGGAATTTCTAGTGTGAACCTGCTGCTTTCTCATCATGTTGTTTC 1701
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 70 TATGGAAGCATGTTTATAGTGTTCATCAAAAGTCCATACAGCACTCAATACGGAAT 129
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1702 TATATTACTATGTCCTGCTCCATTCAAAACCCGCTTGCAGACCAAGAGTAAGGAAT 1761
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 130 CAAGTTAAAAAGAGATGATCTGTCGCAAAAGCTTTTCTTTATAGTATTTACTGATGCA 189
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1762 TGTTTTGAAGAGAGTGGCTGTGCAAAATCGTTTCTTTTATAGTGTTCCTGATGCC 1821
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 190 TTATGCTGGATACCAATTTTGTAGCGAAACCTCTTCACTGCTTCAGGTAGAAATACCA 249
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1822 ATCTGCTGGATTCCTGTATTTAGTATTAATTCCTTCCCTCTCCGGGTGGAATACCA 1881
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 250 GTACCATAACTCTTGGGTAGTATGTTATTCCTCCATTAACAGTGCCTTGAACCCA 309
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

Db 1882 GACACAATGACTTCCTGGATAGTAGTATTTTTCCTCCAGTTAACAGTCGCTTTGAATCCA 1941  
QY 310 ATTCTCTATCTGTGACACACAGACCATTATTAAGAAATGATTCATCGGTTTGGCATAA 368  
Db 1942 ATCCCTATATCTCTACACACCAACTTTTAAAGGACACAGTTGAACAGCTGCTGCACAA 2000

Search completed: September 25, 2003, 15:11:48  
Job time : 191 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 25, 2003, 14:52:51 ; Search time 41 Seconds  
(without alignments)  
607.807 Million cell updates/sec

Title: US-10-049-568-2  
Perfect score: 826  
Sequence: 1 AQIYSVAIFLGINAARPIII.....HMLHSSGGKCGHCRCHLS 157

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: A\_Geneseq\_13Jun03.\*

1: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*  
24: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Query	Score	Match	Length	ID	Description
1	826	100.0	157	22	AA035407	Human HGRL101 G-pr
2	638	77.2	188	22	AA099945	Human expressed po
3	638	77.2	209	22	AA087243	Novel central nerv
4	638	77.2	386	21	AA057286	Human GPCR protein
5	638	77.2	396	22	AA02498	Human GPCR protein
6	638	77.2	396	23	AA054868	Human G-protein co
7	638	77.2	722	20	AA042171	Human LGR7 short f
8	638	77.2	757	20	AA042170	Human LGR7 long fo
9	638	77.2	757	24	AA081724	Human G protein-co

10	632	76.5	176	22	AA087545	Novel central nerv
11	632	76.5	176	22	AA099972	Human expressed po
12	632	76.5	176	22	AA04062	Human musculoskele
13	632	76.5	176	22	AA010249	Human cDNA SEQ ID
14	632	76.5	176	22	AA018108	Novel human uterin
15	632	76.5	176	22	AA018344	Human endocrine po
16	632	76.5	176	22	AA018666	Renal and cardiova
17	632	76.5	176	22	AA021646	Novel human neopla
18	632	76.5	176	23	AA068836	Human polypeptide
19	632	76.5	176	23	AA057335	Novel human protei
20	632	76.5	176	24	AA013356	Novel human muscu
21	586.5	71.0	140	21	AA041526	Human ORFX ORF1290
22	392.5	47.5	355	22	AA04370	Human G-protein co
23	392.5	47.5	355	23	AA08253	Human G-protein-co
24	392.5	47.5	491	23	AA06252	Human G-protein-co
25	392.5	47.5	515	23	AA06251	Human G-protein-co
26	392.5	47.5	610	23	AA06250	Human G-protein-co
27	392.5	47.5	713	23	AA097159	Human G-protein co
28	392.5	47.5	726	23	AA071322	Human GPCR-1 INCY
29	392.5	47.5	730	23	AA06255	Human G-protein-co
30	392.5	47.5	737	23	AA097158	Human G-protein-co
31	392.5	47.5	754	23	AA06254	Human G-protein-co
32	392.5	47.5	754	24	AA016303	Human LGR9 G-prote
33	391.5	47.4	363	24	AA037884	NOVX protein seque
34	351	42.5	108	23	AA020167	Human nGPCR-1079 V
35	349	42.3	107	23	AA020148	Human novel G-prot
36	349	42.3	109	23	AA020168	Human nGPCR-1079 V
37	244	29.5	327	22	AA080961	Human nGPCR57. Ho
38	244	29.5	327	23	AA037779	Human G-protein-co
39	224.5	27.2	334	22	AA060463	Drosophila melanog
40	224.5	27.2	334	22	AA038929	Drosophila G-prote
41	212.5	25.7	359	22	AA061216	Drosophila melanog
42	212.5	25.7	359	22	AA038931	Drosophila G-prote
43	207.5	25.1	1115	23	AA047959	Lymnaea stagnalis
44	139.5	16.9	620	14	AA030522	N-terminal of LH r
45	139.5	16.9	634	14	AA030520	N-terminal of LH r

#### ALIGNMENTS

RESULT 1  
AA035407  
ID AAB35407 standard; Protein; 157 AA.

XX AAB35407;

XX AC

XX DT 23-MAY-2001 (first entry)

XX DE Human HGRL101 G-protein coupled receptor.

XX KW Human; HGRL101; G-protein coupled receptor; infection; pain; cancer;

XX KW diabetes; obesity; eating disorder; asthma; Parkinson's disease;

XX KW hypotension; osteoporosis; myocardial infarction; migraine; allergy;

XX KW psychotic disorder; neurological disorder; dyskinesia; vaccine.

XX OS Homo sapiens.

XX PN WO200114548-A2.

XX PD 01-MAR-2001.

XX PF 09-AUG-2000; 2000WO-EP07723.

XX PR 19-AUG-1999; 99EP-0116345.

XX PA (MERE ) MERCK PATENT GMBH.

XX PI Duecker K;

XX DR WPI; 2001-226617/23.

XX DR N-PSDB; AAF28059.

PT Novel G-coupled protein receptor, HGRL101 useful for treating diseases  
PT such as microbial infections, cancers, obesity, asthma, diabetes,  
PT hypertension, osteoporosis, myocardial infarction, stroke, ulcer,  
PT allergy -  
XX  
PS Claim 1; Page 36; 36pp; English.  
XX  
CC The present invention provides the protein and coding sequences for a  
CC novel human G-protein coupled receptor, designated HGRL101. The sequences  
CC are useful in the diagnosis, prevention and treatment of diseases  
CC including infections, pain, cancer, diabetes, anorexia, bulimia, asthma,  
CC Parkinson's disease, acute heart failure, hypotension, hypertension,  
CC urinary retention, osteoporosis, angina pectoris, myocardial infarction,  
CC stroke, ulcers, allergies, benign prostatic hypertrophy, migraine,  
CC vomiting, psychotic and neurological disorders such as anxiety,  
CC schizophrenia, manic depression, depression, delirium, dementia and  
CC severe mental retardation, and dyskinesias including Huntington's disease  
CC and Gilles de la Tourette's syndrome. The present sequence is the HGRL101  
CC protein.  
XX  
SQ Sequence 157 AA;  
Query Match 100.0%; Score 826; DB 22; Length 157;  
Best Local Similarity 100.0%; Pred. No. 1.3e-88;  
Matches 157; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 AQTYSVAIFGIGNALPFIIVFSGMSFYVSHQSATATEIRNOVKKEMILAKFFFIYF 60  
DB 1 AQTYSVAIFGIGNALPFIIVFSGMSFYVSHQSATATEIRNOVKKEMILAKFFFIYF 60  
QY 61 TDALCWPIFVAKPLSLLOVEIPGTTTSVVVIGVSAINALNPILYTLTTRPKKEMIHFR 120  
DB 61 TDALCWPIFVAKPLSLLOVEIPGTTTSVVVIGVSAINALNPILYTLTTRPKKEMIHFR 120  
QY 121 WHNYRQRKSMDSKGIKRMHLHSSGKCGHCRCHLS 157  
DB 121 WHNYRQRKSMDSKGIKRMHLHSSGKCGHCRCHLS 157  
RESULT 2  
AAM99945  
ID AAM99945 standard; Protein; 188 AA.  
XX  
AC AAM99945;  
XX  
DT 04-JAN-2002 (first entry)  
XX  
DE Human expressed polypeptide SEQ ID NO 69.  
XX  
KW Human; nontropic; neuroprotective; cytostatic; dermatological; virucide;  
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnery;  
KW antiparkinsonian; antiscikling; antianaemic; antiarthritic; cancer;  
KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;  
KW antiallergic; antidiabetic; antiulcer; anticonvulsant; antifungal;  
KW antiparasitic; cardiac; immune disorder; cardiovascular disorder;  
KW neurological disease; infection; nephrotropic; gene therapy; vaccine.  
XX  
OS Homo sapiens.  
XX  
PN WO200155387-A1.  
XX  
PD 02-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US01310.  
XX  
PR 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227189.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.

PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241788.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0257119.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-02534097.  
PR 05-JAN-2001; 2001US-02539678.  
XX  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Barash SC, Ruben SM;  
PI  
XX  
XX WPI: 2001-465573/50.  
DR N-PSDB; AA199557.  
XX  
XX  
XX Isolated digestive system associated polypeptide for treating,  
PT preventing and/or prognosing disorders related to the digestive system  
PT including digestive system cancers and also for testing and detection  
PT e.g. diagnosis -  
XX  
XX Claim 11; SEQ ID NO 69; 509pp + Sequence Listing; English.  
XX  
XX The invention relates to novel genes (AA199548-AA199604) and proteins  
CC (AA199536-AA199584) useful for preventing, treating or ameliorating  
CC medical conditions e.g. by protein or gene therapy. The genes are

CC isolated from a range of human tissues disclosed in the specification.  
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful  
CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast  
CC and ovarian cancer and other cancers of the adrenal gland, bone, bone  
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;  
CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune  
CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's  
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative  
CC colitis; (c) cardiovascular disorders such as myocardial ischaemia;  
CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and  
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal  
CC and parasitic infections.  
CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 188 AA;  
Query Match 77.2%; Score 638; DB 22; Length 188;  
Best Local Similarity 94.8%; Pred. No. 1.8e-66;  
Matches 127; Conservative 1; Mismatches 6; Indels 0; Gaps 0;  
QY 1 AQIYSVAIFLGINLAAPFIIVFSYGMFYSVHQSAITATIRNQVKEMILAKRFFVIF 60  
Db 7 AQIYSVAIFLGINLAAPFIIVFSYGMFYSVHQSAITATIRNQVKEMILAKRFFVIF 66  
QY 61 TDALCWIPFIYAKPLSLLOVEIPGRITTSWVITGYSAINSLNFIILYLTTRPKEMIHFR 120  
Db 67 TDALCWIPFIYVVKELSLLOVEIPGITTSWVIFIPINSALNFILITLTPRKEMIHFR 126  
QY 121 WNYRQRKSDSKG 134  
Db 127 WNYRQRKSDSKG 140  
RESULT 3  
AAU87243  
ID AAU87243 standard; Protein; 209 AA.  
XX  
AC AAU87243;  
XX  
DT 05-JUN-2002 (first entry)  
DE Novel central nervous system protein #153.  
XX  
KW Central nervous system; CNS; autoimmune disease; rheumatoid arthritis;  
KW hyperproliferative disorder; neoplasm; cardiovascular disorder;  
KW cardiac arrest; cerebrovascular disorder; ischaemia; angiogenesis;  
KW nervous system disorder; Alzheimer's disease; AIDS; ocular disorder;  
KW acquired immunodeficiency virus; dysphagia; gastrointestinal disorder;  
KW adenocarcinoma; reproductive system disorder; testicular feminisation;  
KW endocrine disorder; diabetes; cancer; leukaemia; neovascularisation;  
KW respiratory disorder; renal disorder; kidney failure; blood disorder;  
KW myocardial infarction; wound healing; cell proliferation; skin aging;  
KW food additive; food preservative; gene therapy.  
XX  
OS Homo sapiens.  
XX  
PN WO200155318-A2.  
XX  
PD 02-AUG-2001.  
XX  
XX  
PF 17-JAN-2001; 2001WO-US01332.  
XX  
XX 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184684.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.

PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 14-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 18-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 03-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.

PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249267.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.

(HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Barash SC, Ruben SM;

XX WPI; 2001-581633/65.  
XX N-PSDB; ABK43573.

PT New isolated nucleic acid encoding a protein for diagnosing,  
PT preventing, treating or ameliorating medical conditions and used as  
PT food additives or preservatives -

PS Claim 9; SEQ ID No 761; 837pp; English.

XX The invention describes an isolated nucleic acid molecule (I) encoding a  
CC novel central nervous system protein. (I) and polypeptides (III) encoded  
CC by (I), are used to treat a medical conditions and in diagnosis of a  
CC pathological condition. Disorders which are diagnosed or treated include  
CC autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative

CC disorders e.g. neoplasms of the breast or liver, cardiovascular disorders  
 CC e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia,  
 CC angiotensin, nervous system disorders e.g. Alzheimer's disease and  
 CC amyotrophic lateral sclerosis, infections caused by bacteria, viruses  
 CC e.g. Acquired immunodeficiency virus (AIDS) and fungi, ocular disorders  
 CC e.g. corneal infection, gastrointestinal disorders e.g. dysphagia,  
 CC adenocarcinomas and irritable bowel syndrome, reproductive system  
 CC disorders e.g. testicular feminisation, endocrine disorders e.g. diabetes  
 CC and pituitary dwarfism, cancers and disorders at the cellular level e.g.  
 CC leukaemia, disorders involving neovascularisation e.g. malignancies,  
 CC respiratory disorders e.g. nonallergic rhinitis, renal disorders e.g.  
 CC acute kidney failure and blood related disorders e.g. myocardial  
 CC infarction. The polypeptides can also be used to aid wound healing and  
 CC epithelial cell proliferation, to prevent skin aging due to sunburn, to  
 CC maintain organs before transplantation, for supporting cell culture of  
 CC primary tissues, to regenerate tissues and in chemotaxis. The  
 CC polypeptides can also be used as a food additive or preservative to  
 CC increase or decrease storage capabilities, fat content, lipid, protein,

Query Match 77.28; Score 638; DB 22; Length 209;  
 Best Local Similarity 94.8%; Pred. No. 2e-66;  
 Matches 127; Conservative 1; Mismatches 6; Indels 0; Gaps 0;  
 Qy 1 AQIYSVAIFLGINLAAPFIIIVFSYSGMFYSVHQSATATATIRNQVKEMILAKRFFFI 60  
 Db 28 AQIYSVAIFLGINLAAPFIIIVFSYSGMFYSVHQSATATATIRNQVKEMILAKRFFFI 87  
 Qy 61 TDALCWIPFVAKPLSLLOVEIPGTTTSWVIGYSAINSLNPILYLTTRPFKEMIH 120  
 Db 88 TDALCWIPFVAKPLSLLOVEIPGTTTSWVIFILPINSALNPILYLTTRPFKEMIH 147  
 Qy 121 WHNYQRKSMDSKG 134  
 Db 148 WHNYQRKSMDSKG 161

RESULT 4  
 AAY57286  
 ID AAY57286 standard; Protein; 396 AA.  
 AC AAY57286;  
 XX 05-JUN-2000 (first entry)  
 DT Human GPCR protein (HGPRP) sequence (clone ID 2488822).  
 DE Human; G protein coupled protein receptor; HGPRP; cell proliferation;  
 KW neurological; immune disorder; cytostatic; anti-arteriosclerotic;  
 KW anti-atherosclerotic; hepatotropic; antiinflammatory; virucide; leukemia;  
 KW immunomodulatory; anemia; asthma; gastrointestinal; anti-epileptic;  
 KW anti-Alzheimer's; anti-Parkinsonian; gene therapy.  
 XX Homo sapiens.  
 OS WO200015793-A2.  
 PN 23-MAR-2000.  
 XX 17-SEP-1999; 99WO-US20958.  
 XX 17-SEP-1998; 98US-0156513.  
 PR (INCYTE) INCYTE PHARM INC.  
 XX Bandman O, Lal P, Tang YT, Corley NC, Guegler KJ, Gorgone GA;  
 PI Baughn MR;  
 XX WPI; 2000-271432/23.  
 DR N-PSDB; AAZ90524.  
 XX Human G protein coupled protein receptor peptides useful for the  
 PT prevention, diagnosis and treatment of cell proliferative, neurological  
 PT and immune disorders -

XX Claim 1; Page 62-63; 7lpp; English.  
 PS The invention provides human G protein coupled protein receptor (HGPRP)  
 XX polypeptides and polynucleotides encoding them. The polypeptides can be  
 CC produced by standard recombinant methodology. The polynucleotides and  
 CC polypeptides may be used in the prevention, treatment and diagnosis of  
 CC diseases associated with their inappropriate expression. Diseases that  
 CC can be treated are cell proliferative disorders (e.g. arteriosclerosis,  
 CC atherosclerosis and hepatitis), cancers (e.g. leukemia, melanomas and  
 CC adenocarcinoma), immune disorders (e.g. anemia, asthma and Crohn's  
 CC disease) and neurological disorders (e.g. epilepsy, Alzheimer's disease  
 CC and Parkinson's disease). The anti-HGPRP antibodies may also be used as  
 CC diagnostic agents for detecting the presence of HGPRP polypeptides in  
 CC samples (e.g. by enzyme linked immunosorbent assay (ELISA)). Sequences  
 CC AAY57283-288 represent the HGPRP polypeptides.  
 XX Sequence 396 AA;

Query Match 77.28; Score 638; DB 21; Length 396;  
 Best Local Similarity 94.8%; Pred. No. 4.7e-66;  
 Matches 127; Conservative 1; Mismatches 6; Indels 0; Gaps 0;  
 Qy 1 AQIYSVAIFLGINLAAPFIIIVFSYSGMFYSVHQSATATATIRNQVKEMILAKRFFFI 60  
 Db 215 AQIYSVAIFLGINLAAPFIIIVFSYSGMFYSVHQSATATATIRNQVKEMILAKRFFFI 274  
 Qy 61 TDALCWIPFVAKPLSLLOVEIPGTTTSWVIGYSAINSLNPILYLTTRPFKEMIH 120  
 Db 275 TDALCWIPFVAKPLSLLOVEIPGTTTSWVIFILPINSALNPILYLTTRPFKEMIH 334  
 Qy 121 WHNYQRKSMDSKG 134  
 Db 335 WHNYQRKSMDSKG 348

RESULT 5  
 AAE02498  
 ID AAE02498 standard; Protein; 396 AA.  
 AC AAE02498;  
 XX 10-AUG-2001 (first entry)  
 DT Human CON222 G protein-coupled receptor protein.  
 DE Human; G protein-coupled receptor; GPCR; CON222 protein; schizophrenia;  
 KW neuroleptic; nootropic; neuroprotective; bipolar disease; psychotropic;  
 KW neurological disorder; psychiatric disease; neurosis; anxiety; dementia;  
 KW attention deficit hyperactivity disorder; neurasthenia; senile dementia;  
 KW affective disorder; neuropathy; Alzheimer's disease; Parkinson's disease;  
 KW depression; migraine; genetic screening.  
 XX Homo sapiens.  
 OS Location/Qualifiers  
 XX Key 42..65  
 FH Domain /label= Transmembrane\_domain\_(1TM)  
 FT 66..78  
 FT Domain /label= Intracellular\_domain  
 FT /note= "First IC loop"  
 FT 79..103  
 FT Domain /label= Transmembrane\_domain\_(2TM)  
 FT 104..124  
 FT Domain /label= Extracellular\_domain  
 FT /note= "First EC loop"  
 FT 125..156  
 FT Domain /label= Transmembrane\_domain\_(3TM)  
 FT 157..166  
 FT Domain /label= Intracellular\_domain  
 FT /note= "Second IC loop"  
 FT 167..188  
 FT Domain /label= Transmembrane\_domain\_(4TM)

Domain 189..216  
/label= Extracellular\_domain  
/note= "Second EC loop"  
217..241  
Domain  
/label= Transmembrane\_domain\_(5TM)  
242..267  
Domain  
/note= Intracellular\_domain  
/label= "Third IC loop"  
268..290  
Domain  
/label= Transmembrane\_domain\_(6TM)  
291..300  
Domain  
/label= Extracellular\_domain  
/note= "Third EC loop"  
301..320  
Domain  
/label= Transmembrane\_domain\_(7TM)  
WO200131014-A2.  
03-MAY-2001.  
27-OCT-2000; 2000WO-US29601.  
27-OCT-1999; 99US-0427653.  
27-OCT-1999; 99US-0427859.  
27-OCT-1999; 99US-0428020.  
27-OCT-1999; 99US-0428114.  
28-OCT-1999; 99US-0429517.  
28-OCT-1999; 99US-0429555.  
28-OCT-1999; 99US-0429675.  
28-OCT-1999; 99US-0429695.  
03-DEC-1999; 99US-0454399.  
12-JAN-2000; 2000US-0481794.  
(PHAA ) PHARMACIA & UPJOHN CO.  
Vogeli G, Wood LS, Merchant K;  
WPI; 2001-328653/34.  
N-PSDB; RAD06507.  
Seven transmembrane receptor polypeptides and polynucleotides, useful  
for treating neurological or psychiatric disorders, e.g. schizophrenia,  
as well as for identifying compounds useful for treating schizophrenia  
Claim 1; Page 15-16; 215pp; English.  
The invention relates to human G protein-coupled receptor (GPCR) and  
their corresponding DNA molecules. GPCR is also referred as seven  
transmembrane receptor. G protein-coupled receptor protein is useful for  
treating neurological disorder, particularly schizophrenia. GPCR protein  
is also useful for identifying compounds useful for treating other  
schizophrenia. These compounds are also useful for treating other  
neurological and psychiatric diseases, e.g. depression, anxiety, bipolar  
disease, affective disorders, attention deficit hyperactivity disorder/  
attention deficit disorder, epilepsy, neuritis, neuroasthenia, neuropathy,  
neurosis, Alzheimer's disease, Parkinson's disease, migraine and senile  
dementia. The invention also provides genetic screening procedures that  
entail analysing a person's genome with respect to GPCR. The vectors are  
useful for the recombinant production of the GPCR's. The present sequence  
is human CON222 G protein-coupled receptor (GPCR) protein.  
Sequence 396 AA;  
Query Match 77.2%; Score 638; DB 22; Length 396;  
Best Local Similarity 94.8%; Pred. No. 4.7e-66;  
Matches 127; Conservative 1; Mismatches 6; Indels 0; Gaps 0;  
1 AQIYSVAIFGINLAFIIVFSYSGMFYSVHOSAITATEIRNQVKEMILAKRFFFIYF 60  
215 AQIYSVAIFGINLAFIIVFSYSGMFYSVHOSAITATEIRNQVKEMILAKRFFFIYF 274  
61 TDALCWIPFVAKPFLSLQVEIPGTTTSWVVGYSAINSLNPILYITLTRPFKEMIHFRF 120

Db 275 TDALCWIPFVVKFSLQVEIPGTTTSWVVGYSAINSLNPILYITLTRPFKEMIHFRF 334  
QY 121 WHNYORKSMDSKG 134  
Db 335 WYNYORKSMDSKG 348  
RESULT 6  
ABU54868  
ID ABU54868 standard; Protein; 396 AA.  
XX  
AC ABU54868;  
XX  
DT 14-MAR-2003 (first entry)  
XX  
DE Human G-protein coupled receptor #4, Incyte clone 2488822CD1.  
XX  
KW Human; G-protein coupled receptor; GPCR; neoplastic disorder;  
KW neurological disorder; immune disorder; cytostatic; pancreatic cancer;  
KW follicular carcinoma of the thyroid; leiomyoma of the uterus; epilepsy;  
KW interstitial nephritis; cancer.  
XX  
OS Homo sapiens.  
XX  
PN US2002106655-A1.  
XX  
PD 08-AUG-2002.  
XX  
PF 28-JUN-2001; 2001US-0895686.  
XX  
PR 17-SEP-1998; 98US-0156513.  
XX  
PA (BAND/) BANDMAN O.  
PA (LALP/) LAL P G.  
PA (TANG/) TANG Y T.  
PA (BAUG/) BAUGHN M R.  
XX  
PI Bandman O, Lal PG, Tang YT, Baughn MR;  
XX  
DR WPI; 2002-697866/75.  
DR N-PSDB; ABX73054.  
XX  
PT New cDNAs encoding G protein coupled receptors are useful for the  
PT diagnosis, prognosis, treatment and evaluation of therapies for  
PT neoplastic, neurological and immune disorders  
XX  
PS Claim 13; Fig 2; 61pp; English.  
XX  
CC The invention relates to an isolated cDNA encoding G-protein  
CC coupled receptor (GPCR) appearing as ABU54865-ABU54870.  
CC Also included are fragments of the cDNAs, species variants having  
CC at least 75% identity to the cDNAs, vectors comprising the cDNAs,  
CC a host cells comprising the above vectors, producing a protein  
CC (comprising culturing the above host cell under expression conditions and  
CC recovering the protein), using a cDNA to detect expression of a  
CC nucleic acid in a sample or to screen for compounds or molecules  
CC which bind to the cDNAs, using the GPCR proteins to screen  
CC compounds or molecules for ligands, using a GPCR protein to prepare and  
CC purify antibodies, an anti-GPCR antibody and using the antibody to detect  
CC expression of a GPCR protein in a sample and is diagnostic of cancer.  
CC The invention is useful for the diagnosis, prognosis, treatment  
CC and evaluation of therapies for neoplastic, neurological and immune  
CC disorders, particularly follicular carcinoma of the thyroid, leiomyoma  
CC of the uterus, pancreatic cancer, epilepsy, interstitial nephritis and  
CC immune response as a complication of cancer. The present sequence  
CC is a human GPCR protein of the invention.  
XX  
SQ Sequence 396 AA;  
Query Match 77.2%; Score 638; DB 23; Length 396;  
Best Local Similarity 94.8%; Pred. No. 4.7e-66;  
Matches 127; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

```

QY 1 AQIYSVAIFLGINLAAPITIIIVFSYGMFYSVHQSAITATEIRNOVKKEMILAKRFFFIYF 60
DB 215 AQIYSVAIFLGINLAAPITIIIVFSYGMFYSVHQSAITATEIRNOVKKEMILAKRFFFIYF 274

QY 61 TDALCWIPFVAKPLSLLOVEIPGTTISWVIGYSAINSLNPILYTLTTRPFKEMIHFR 120
DB 275 TDALCWIPFVAKPLSLLOVEIPGTTISWVIFILPINSALNPILYTLTTRPFKEMIHFR 334

QY 121 WHNYRQKSMDSKG 134
DB 335 WYNYRQKSMDSKG 348

RESULT 7
RAY42171
ID AAY42171 standard; Protein; 722 AA.
XX
AC AAY42171;
XX
DT 20-DEC-1999 (first entry)
XX
DE Human LGR7 short form protein sequence.
XX
KW Human; LGR4; LGR5; LGR7; G-protein coupled receptor; gene therapy;
KW extracellular leucine rich repeat region; mapping; identification.
XX
OS Homo sapiens.
XX
PN WO9948921-A1.
XX
PD 30-SEP-1999.
XX
PF 25-MAR-1999; 99WO-US06573.
XX
PR 26-MAR-1998; 98US-0079501.
XX
PA (STRD ) UNIV LELAND STANFORD JUNIOR.
PA (ORGA ) ORGANON NV.
XX
PI Hsueh AJW, Hsu SY, Liang S, Van Der Spek PJ;
XX
DR WPI; 1999-591074/50.
DR N-PSDB; AA225346.
XX
PT New G-protein coupled receptors, useful for identifying their own
ligands -
XX
PS Claim 2; Fig 4; 54pp; English.
XX
CC The present sequence represents the human G-protein coupled receptor
CC having extracellular leucine rich repeat regions, designated LGR7 short
CC form. The LGR4, LGR5 and LGR7 proteins are used to identify ligands for
CC the receptor. The polypeptides and/or polynucleotides are also useful
CC for homologous or related genes, producing compositions that modulate
CC the expression or function of the receptors, gene therapy, mapping
CC functional regions of the receptors, studying associated physiological
CC pathways, in vivo prophylactic and therapeutic purposes, as immunogens
CC for producing antibodies, and for identifying biologically active
CC agents. The polypeptides contain a G-protein coupled seven
CC transmembrane region and a leucine rich repeat extracellular domain.
CC These regions capture and facilitate optimal orientation of its ligand.
CC The proteins are also expressed in diverse tissues.
XX
SQ Sequence 722 AA;

Query Match 77.2%; Score 638; DB 20; Length 722;
Best Local Similarity 94.8%; Pred. No. 1e-65;
Matches 127; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 AQIYSVAIFLGINLAAPITIIIVFSYGMFYSVHQSAITATEIRNOVKKEMILAKRFFFIYF 60
DB 541 AQIYSVAIFLGINLAAPITIIIVFSYGMFYSVHQSAITATEIRNOVKKEMILAKRFFFIYF 600

QY 1 AQIYSVAIFLGINLAAPITIIIVFSYGMFYSVHQSAITATEIRNOVKKEMILAKRFFFIYF 60
DB 576 AQIYSVAIFLGINLAAPITIIIVFSYGMFYSVHQSAITATEIRNOVKKEMILAKRFFFIYF 635

QY 61 TDALCWIPFVAKPLSLLOVEIPGTTISWVIGYSAINSLNPILYTLTTRPFKEMIHFR 120
DB 636 TDALCWIPFVAKPLSLLOVEIPGTTISWVIFILPINSALNPILYTLTTRPFKEMIHFR 695

```

```

QY 61 TDALCWIPFVAKPLSLLOVEIPGTTISWVIGYSAINSLNPILYTLTTRPFKEMIHFR 120
DB 601 TDALCWIPFVAKPLSLLOVEIPGTTISWVIFILPINSALNPILYTLTTRPFKEMIHFR 660

QY 121 WHNYRQKSMDSKG 134
DB 661 WYNYRQKSMDSKG 674

RESULT 8
RAY42170
ID AAY42170 standard; Protein; 757 AA.
XX
AC AAY42170;
XX
DT 20-DEC-1999 (first entry)
XX
DE Human LGR7 long form protein sequence.
XX
KW Human; LGR4; LGR5; LGR7; G-protein coupled receptor; gene therapy;
KW extracellular leucine rich repeat region; mapping; identification.
XX
OS Homo sapiens.
XX
PN WO9948921-A1.
XX
PD 30-SEP-1999.
XX
PF 25-MAR-1999; 99WO-US06573.
XX
PR 26-MAR-1998; 98US-0079501.
XX
PA (STRD ) UNIV LELAND STANFORD JUNIOR.
PA (ORGA ) ORGANON NV.
XX
PI Hsueh AJW, Hsu SY, Liang S, Van Der Spek PJ;
XX
DR WPI; 1999-591074/50.
DR N-PSDB; AA225345.
XX
PT New G-protein coupled receptors, useful for identifying their own
ligands -
XX
PS Claim 2; Fig 3; 54pp; English.
XX
CC The present sequence represents the human G-protein coupled receptor
CC having extracellular leucine rich repeat regions, designated LGR7 long
CC form. The LGR4, LGR5 and LGR7 proteins are used to identify ligands for
CC the receptor. The polypeptides and/or polynucleotides are also useful
CC for homologous or related genes, producing compositions that modulate
CC the expression or function of the receptors, gene therapy, mapping
CC functional regions of the receptors, studying associated physiological
CC pathways, in vivo prophylactic and therapeutic purposes, as immunogens
CC for producing antibodies, and for identifying biologically active
CC agents. The polypeptides contain a G-protein coupled seven
CC transmembrane region and a leucine rich repeat extracellular domain.
CC These regions capture and facilitate optimal orientation of its ligand.
CC The proteins are also expressed in diverse tissues.
XX
SQ Sequence 757 AA;

Query Match 77.2%; Score 638; DB 20; Length 757;
Best Local Similarity 94.8%; Pred. No. 1.1e-65;
Matches 127; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 AQIYSVAIFLGINLAAPITIIIVFSYGMFYSVHQSAITATEIRNOVKKEMILAKRFFFIYF 60
DB 576 AQIYSVAIFLGINLAAPITIIIVFSYGMFYSVHQSAITATEIRNOVKKEMILAKRFFFIYF 635

QY 61 TDALCWIPFVAKPLSLLOVEIPGTTISWVIGYSAINSLNPILYTLTTRPFKEMIHFR 120
DB 636 TDALCWIPFVAKPLSLLOVEIPGTTISWVIFILPINSALNPILYTLTTRPFKEMIHFR 695

```





PR 14-AUG-2000; 2000US-0225213.  
 PR 14-AUG-2000; 2000US-0225214.  
 PR 14-AUG-2000; 2000US-0225266.  
 PR 14-AUG-2000; 2000US-0225267.  
 PR 14-AUG-2000; 2000US-0225268.  
 PR 14-AUG-2000; 2000US-0225270.  
 PR 14-AUG-2000; 2000US-0225447.  
 PR 14-AUG-2000; 2000US-0225757.  
 PR 14-AUG-2000; 2000US-0225758.  
 PR 14-AUG-2000; 2000US-0225759.  
 PR 18-AUG-2000; 2000US-0226279.  
 PR 22-AUG-2000; 2000US-0226681.  
 PR 22-AUG-2000; 2000US-0226686.  
 PR 22-AUG-2000; 2000US-0227182.  
 PR 23-AUG-2000; 2000US-0227009.  
 PR 30-AUG-2000; 2000US-0228924.  
 PR 01-SEP-2000; 2000US-0229287.  
 PR 01-SEP-2000; 2000US-0229343.  
 PR 01-SEP-2000; 2000US-0229344.  
 PR 01-SEP-2000; 2000US-0229345.  
 PR 05-SEP-2000; 2000US-0229509.  
 PR 05-SEP-2000; 2000US-0229513.  
 PR 06-SEP-2000; 2000US-0230437.  
 PR 06-SEP-2000; 2000US-0230438.  
 PR 08-SEP-2000; 2000US-0231242.  
 PR 08-SEP-2000; 2000US-0231243.  
 PR 08-SEP-2000; 2000US-0231244.  
 PR 08-SEP-2000; 2000US-0231413.  
 PR 08-SEP-2000; 2000US-0231414.  
 PR 08-SEP-2000; 2000US-0232080.  
 PR 08-SEP-2000; 2000US-0232081.  
 PR 12-SEP-2000; 2000US-0231968.  
 PR 14-SEP-2000; 2000US-0232397.  
 PR 14-SEP-2000; 2000US-0232398.  
 PR 14-SEP-2000; 2000US-0232399.  
 PR 14-SEP-2000; 2000US-0232400.  
 PR 14-SEP-2000; 2000US-0232401.  
 PR 14-SEP-2000; 2000US-0233063.  
 PR 14-SEP-2000; 2000US-0233064.  
 PR 14-SEP-2000; 2000US-0233065.  
 PR 21-SEP-2000; 2000US-0234223.  
 PR 21-SEP-2000; 2000US-0234274.  
 PR 25-SEP-2000; 2000US-0234997.  
 PR 25-SEP-2000; 2000US-0234998.  
 PR 26-SEP-2000; 2000US-0235484.  
 PR 27-SEP-2000; 2000US-0235834.  
 PR 27-SEP-2000; 2000US-0235836.  
 PR 29-SEP-2000; 2000US-0236327.  
 PR 29-SEP-2000; 2000US-0236367.  
 PR 29-SEP-2000; 2000US-0236368.  
 PR 29-SEP-2000; 2000US-0236369.  
 PR 29-SEP-2000; 2000US-0236370.  
 PR 02-OCT-2000; 2000US-0236802.  
 PR 02-OCT-2000; 2000US-0237037.  
 PR 02-OCT-2000; 2000US-0237038.  
 PR 02-OCT-2000; 2000US-0237039.  
 PR 02-OCT-2000; 2000US-0237040.  
 PR 13-OCT-2000; 2000US-0239935.  
 PR 13-OCT-2000; 2000US-0239937.  
 PR 20-OCT-2000; 2000US-0240960.  
 PR 20-OCT-2000; 2000US-0241221.  
 PR 20-OCT-2000; 2000US-0241785.  
 PR 20-OCT-2000; 2000US-0241786.  
 PR 20-OCT-2000; 2000US-0241787.  
 PR 20-OCT-2000; 2000US-0241808.  
 PR 20-OCT-2000; 2000US-0241809.  
 PR 20-OCT-2000; 2000US-0241826.  
 PR 01-NOV-2000; 2000US-0244617.  
 PR 08-NOV-2000; 2000US-0246474.  
 PR 08-NOV-2000; 2000US-0246475.  
 PR 08-NOV-2000; 2000US-0246476.  
 PR 08-NOV-2000; 2000US-0246477.  
 PR 08-NOV-2000; 2000US-0246478.

PR 08-NOV-2000; 2000US-0246523.  
 PR 08-NOV-2000; 2000US-0246524.  
 PR 08-NOV-2000; 2000US-0246525.  
 PR 08-NOV-2000; 2000US-0246526.  
 PR 08-NOV-2000; 2000US-0246527.  
 PR 08-NOV-2000; 2000US-0246528.  
 PR 08-NOV-2000; 2000US-0246532.  
 PR 08-NOV-2000; 2000US-0246609.  
 PR 08-NOV-2000; 2000US-0246610.  
 PR 08-NOV-2000; 2000US-0246611.  
 PR 08-NOV-2000; 2000US-0246613.  
 PR 17-NOV-2000; 2000US-0249207.  
 PR 17-NOV-2000; 2000US-0249208.  
 PR 17-NOV-2000; 2000US-0249209.  
 PR 17-NOV-2000; 2000US-0249210.  
 PR 17-NOV-2000; 2000US-0249211.  
 PR 17-NOV-2000; 2000US-0249212.  
 PR 17-NOV-2000; 2000US-0249213.  
 PR 17-NOV-2000; 2000US-0249214.  
 PR 17-NOV-2000; 2000US-0249215.  
 PR 17-NOV-2000; 2000US-0249216.  
 PR 17-NOV-2000; 2000US-0249217.  
 PR 17-NOV-2000; 2000US-0249218.  
 PR 17-NOV-2000; 2000US-0249244.  
 PR 17-NOV-2000; 2000US-0249245.  
 PR 17-NOV-2000; 2000US-0249264.  
 PR 17-NOV-2000; 2000US-0249265.  
 PR 17-NOV-2000; 2000US-0249297.  
 PR 17-NOV-2000; 2000US-0249299.  
 PR 17-NOV-2000; 2000US-0249300.  
 PR 01-DEC-2000; 2000US-0250160.  
 PR 01-DEC-2000; 2000US-0250391.  
 PR 05-DEC-2000; 2000US-0251030.  
 PR 05-DEC-2000; 2000US-0251988.  
 PR 05-DEC-2000; 2000US-0256719.  
 PR 06-DEC-2000; 2000US-0251479.  
 PR 08-DEC-2000; 2000US-0251856.  
 PR 08-DEC-2000; 2000US-0251868.  
 PR 08-DEC-2000; 2000US-0251869.  
 PR 08-DEC-2000; 2000US-0251989.  
 PR 08-DEC-2000; 2000US-0251990.  
 PR 11-DEC-2000; 2000US-0254097.  
 PR 05-JAN-2001; 2001US-0259678.  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 XX Rosen CA, Barash SC, Ruben SM;  
 PI WPI; 2001-581633/65.  
 DR N-PSDB; ABK43875.  
 XX  
 XX New isolated nucleic acid encoding a protein for diagnosing,  
 PT preventing, treating or ameliorating medical conditions and used as  
 PT food additives or preservatives -  
 XX  
 PS Claim 9; SEQ ID No 1063; 837pp; English.  
 XX  
 XX The invention describes an isolated nucleic acid molecule (I) encoding a  
 CC novel central nervous system protein. (I) and polypeptides (III) encoded  
 CC by (I), are used to treat a medical conditions and in diagnosis of a  
 CC pathological condition. Disorders which are diagnosed or treated include  
 CC autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative  
 CC disorders e.g. neoplasms of the breast or liver, cardiovascular disorders  
 CC e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia,  
 CC angiogenesis, nervous system disorders e.g. Alzheimer's disease and  
 CC amyotrophic lateral sclerosis, infections caused by bacteria, viruses  
 CC e.g. Acquired immunodeficiency virus (AIDS) and fungi, ocular disorders  
 CC e.g. corneal infection, gastrointestinal disorders e.g. dysphagia,  
 CC adenocarcinomas and irritable bowel syndrome, reproductive system  
 CC disorders e.g. testicular feminisation, endocrine disorders e.g. diabetes  
 CC and pituitary dwarfism, cancers and disorders at the cellular level e.g.  
 CC leukemia, disorders involving neovascularisation e.g. malignancies,  
 CC respiratory disorders e.g. nonallergic rhinitis, renal disorders e.g.

CC acute kidney failure and blood related disorders e.g. myocardial  
 CC infarction. The polypeptides can also be used to aid wound healing and  
 CC epithelial cell proliferation, to prevent skin aging due to sunburn, to  
 CC maintain organs before transplantation, for supporting cell culture of  
 CC primary tissues, to regenerate tissues and in chemotaxis. The  
 CC polypeptides can also be used as a food additive or preservative to  
 CC increase or decrease storage capabilities, fat content, lipid, protein,

Query Match 76.5%; Score 532; DB 22; Length 176;  
 Best local similarity 94.0%; Pred. No. 8.3e-66;  
 Matches 126; Conservative 1; Mismatches 7; Indels 0; Gaps 0;  
 QY 1 AQIYSVAIFGINLAFAIIIVFSYSGSMFYSHQSAITATEIRNOVKKEMILAKRFFFIYF 60  
 Db 1 AQIYSVAIFGINLAFAIIIVFSYSGSMFYSHQSAITATEIRNOVKKEMILAKRFFFIYF 60  
 QY 51 TDALCWPIFVAKPSLLOVEIPGTITSWVVIYGYSAINSALNPILYLTTRPKEMIHRF 120  
 Db 61 TDALCWPIFVAKPSLLOVEIPGTITSWVVIYGYSAINSALNPILYLTTRPKEMIHRF 120  
 QY 121 WNYRQKMSDVG 134  
 Db 121 WNYRQKMSDVG 134

RESULT 11  
 AAM99972  
 ID AAM99972 standard; Protein; 176 AA.

AC AAM99972;  
 XX  
 DT 04-JAN-2002 (first entry)  
 XX Human expressed polypeptide SEQ ID NO 96.  
 XX Human; neurotropic; neuroprotective; cytostatic; dermatological; virucide;  
 KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;  
 KW antiparkinsonian; antisticking; antianemic; antiarthritic; cancer;  
 KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;  
 KW antiallergic; antidiabetic; antialcifer; anticonvulsant; antifungal;  
 KW antiparasitic; cardiac; immune disorder; cardiovascular disorder;  
 KW neurological disease; infection; nephrotropic; gene therapy; vaccine.

XX Homo sapiens.

OS  
 XX  
 PN WO20015387-A1.

XX  
 XX  
 XX  
 PD 02-AUG-2001.

XX  
 PF 17-JAN-2001; 2001WO-US01310.

XX  
 PR 31-JAN-2000; 2000US-0179065.

PR 04-FEB-2000; 2000US-0180628.

PR 24-FEB-2000; 2000US-0184664.

PR 02-MAR-2000; 2000US-0186350.

PR 16-MAR-2000; 2000US-0189874.

PR 17-MAR-2000; 2000US-0190076.

PR 18-APR-2000; 2000US-0193123.

PR 19-MAY-2000; 2000US-0205515.

PR 07-JUN-2000; 2000US-0209467.

PR 28-JUN-2000; 2000US-0214886.

PR 30-JUN-2000; 2000US-0215135.

PR 07-JUL-2000; 2000US-0216647.

PR 07-JUL-2000; 2000US-0216880.

PR 11-JUL-2000; 2000US-0217487.

PR 11-JUL-2000; 2000US-0217496.

PR 14-JUL-2000; 2000US-0218290.

PR 26-JUL-2000; 2000US-0220963.

PR 26-JUL-2000; 2000US-0220964.

PR 14-AUG-2000; 2000US-0224518.

PR 14-AUG-2000; 2000US-0224519.

PR 14-AUG-2000; 2000US-0225213.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14

[illegible]

CC from WIFO at ftp.wifo.int/pub/published\_pct\_sequences.

XX Sequence 176 AA;

SQ

Query Match 76.58; Score 632; DB 22; Length 176;  
Best Local Similarity 94.08; Pred. NO. 8.3e-66;  
Matches 126; Conservative 1; Mismatches 7; Indels 0; Gaps 0

QY 1 AQIYSVAIFLGINLAFLIIIVFSGSMFYSHQSALTATEIRNOVKKEMILAKRFFFI 60  
Db 1 AQIYSVAIFLGINLAFLIIIVFSGSMFYSHQSALTATEIRNOVKKEMILAKRFFFI 60																															
QY 61 TDALCWIPFVAKPILLSLQVEIPGTISWVVGYSAINSLNPILYTLTRPFKEMIHRF 120																															
Db 61 TDALCWIPFVAKPILLSLQVEIPGTISWVVGYSAINSLNPILYTLTRPFKEMIHRF 120																															
QY 121 WHNYRKSDSKG 134																															
:																															
:																															
Db 121 WHNYRKSDSXG 134

RESULT 12

ABSO4062

ID ID ABBO4062 standard; Protein: 176 AA.

AC ABB04062;

XX

DT 08-JAN-2002 (first entry)

XX

DE Human musculoskeletal system related polypeptide SEQ ID NO 2009.

XX

KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;  
KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;  
KW vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic;  
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;  
KW neurological disease; infection; human; secreted protein;  
KW musculoskeletal system.

XX

XX Homo sapiens.

OS

PN WO200155367-A1.

XX

PD 02-AUG-2001.

XX

PX 17-JAN-2001; 2001WO-US01338.

XX

PR 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184564.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209457.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.

PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0232198.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244517.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250180.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 11-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX Rosen CA, Barash SC, Ruben SM;  
PI WPI: 2001-451937/48.  
XX N-PSDB; AAL35644.  
DR Isolated polypeptide for treating, preventing and/or prognosing  
DR disorders related to the musculoskeletal system including  
DR musculoskeletal cancers and also for testing and detection e.g.  
XX diagnosis  
PS Claim 11; SEQ ID NO 2009; 781pp + Sequence Listing; English.  
XX The invention relates to novel genes (AAL34669-AAL37666) and proteins  
CC (ABB03087-ABB04109) associated with the musculoskeletal system useful  
CC for preventing, treating or ameliorating medical conditions e.g. by  
CC protein or gene therapy. The genes are isolated from a range of human  
CC tissues disclosed in the specification. The nucleic acids, proteins,  
CC antibodies and (ant)agonists are useful in the diagnosis, treatment  
CC and prevention of: (a) cancer, e.g. breast and ovarian cancer and  
CC other cancers of the adrenal gland, bone, bone marrow, breast,  
CC gastrointestinal tract, liver, lung, or urogenital; (b) immune  
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic  
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,  
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis;  
CC (c) cardiovascular disorders such as myocardial ischaemia; (d) wound  
CC healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy;  
CC and (f) infectious diseases such as viral, bacterial, fungal and  
CC parasitic infections.  
CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
XX Sequence 176 AA;  
SQ





PR	08-DEC-2000;	2000US-0251856.
PR	08-DEC-2000;	2000US-0251868.
PR	08-DEC-2000;	2000US-0251869.
PR	08-DEC-2000;	2000US-0251869.
PR	08-DEC-2000;	2000US-0251939.
PR	08-DEC-2000;	2000US-0251980.
PR	11-DEC-2000;	2000US-0254097.
PR	05-JAN-2001;	2001US-0259678.
XX	(HUMA-) HUMAN GENOME SCI INC.	
PA	Rosen CA, Barash SC, Ruben SM;	
PI	WPI; 2001-488777/53.	
DR	N-PDSB; AAS28950.	
XX		
PT	Isolated polypeptide and nucleic acid molecules for treating,	
PT	preventing and/or prognosing disorders related to uterine motility	
PT	e.g. disorders associated with pregnancy and the menstrual cycle -	
PS	Claim 11; SEQ ID NO 84; 524pp; English.	
XX		
CC	The present invention relates to the isolation of novel human	
CC	uterine motility-association polypeptides, and cDNA (AAS28936-AAS28994)	
CC	and genomic sequences encoding for these polypeptides.	
CC	The sequences of the invention are useful in the diagnosis,	
CC	treatment, prevention and/or prognosis of diseases associated	
CC	with uterine motility such as pregnancy and labour, and menstrual	
CC	disorders. The polynucleotide sequences of the invention are also	
CC	useful in gene therapy. AAU18094-AAU18152 represent novel human	
CC	uterine motility-association polypeptides.	
CC	Note: The sequence data for this patent did not form part of the printed	
CC	specification, but was obtained in electronic format directly from WIPO	
CC	at ftp.wipo.int/pub/published_pct_sequences.	
XX		
SQ	Sequence 176 AA;	
	Query Match 76.5%; Score 632; DB 22; Length 176;	
	Best Local Similarity 94.0%; Pred. No. 8.3e-66;	
	Matches 126; Conservative 1; Mismatches 7; Indels 0; Gaps 0	
Qy	1 AQIYSVAIFLGINLAFLIIIVESGMSFYSVHOSATTEIRNOVKKEMILAKRFFEIFV 60	
Db	1 AQIYSVAIFLGINLAFLIIIVESGMSFYSVHOSATTEIRNOVKKEMILAKRFFEIFV 60	
Qy	61 TDALCWIPITFVAKPLSLQLQVEIPGTITSWWYIGYGAINSNLPILTLTRPFKEMHRF 120	
Db	61 TDALCWIPITFVAKPLSLQLQVEIPGTITSWWYIFILPINSALNPILTLTLTRPFKEMHRF 120	
Qy	121 WHNYQRKSMDSKG 134	
Db	121 WYNRYQRKSMDSXG 134	
RESULT 15		
AAU18344		
ID	AAU18344 standard; Protein; 176 AA.	
XX	AAU18344;	
XX		
DT	21-NOV-2001 (first entry)	
XX		
DE	Human endocrine polypeptide SEQ ID NO 299.	
KW	Endocrine protein; human; mouse; rabbit; goat; horse; food additive;	
KW	cat; dog; chicken; sheep; immunosuppressive; antiarthritic; vasotropic;	
KW	antirheumatic; antiproliferative; cytostatic; cardiant; neuroprotective;	
KW	cerebroprotective; nootropic; antibacterial; virucide; fungicide; cancer;	
KW	ophthalmologicall; vulnerary; gene therapy; autoimmune disease; neoplasia;	
KW	hyperproliferative disorder; breast; liver; cardiovascular disorder;	
KW	cerebrovascular disorder; nervous system disorder; bacterial infection;	
KW	fungal infection; viral infection; ocular disorder; endocrine disorder;	
KW	gastrointestinal disorder; renal disorder; respiratory disorder;	
KW	wound healing; skin aging; organ transplantation; food preservative;	





XX (HUMA-) HUMAN GENOME SCI INC.  
XX PA  
XX PI  
XX PI Rosen CA, Barash SC, Ruben SM;  
XX WPI; 2001-451936/48.  
DR N-PSDB; AAS29573.  
XX  
XX Isolated polypeptide for treating, preventing and/or prognosing  
PT disorders of the endocrine system such as reproductive disorders,  
PT endocrine cancers and also for testing and detection e.g. diagnosis -  
XX  
XX Claim 11; SEQ ID No 299; 604pp; English.  
XX  
XX Sequences AAU18282-AAU18507 represent endocrine polypeptides of the  
CC invention. Endocrine polypeptides and their associated polynucleotides  
CC are useful in the diagnosis, treatment and prevention of various types of  
CC disorders in e.g. humans, mice, rabbits, goats, horses, cats, dogs,  
CC chickens or sheep. A pathological condition can be determined by  
CC determining the presence or absence of a mutation in an endocrine  
CC polynucleotide. The treatable disorders include autoimmune diseases such  
CC as rheumatoid arthritis, hyperproliferative disorders such as neoplasms  
CC of the breast or liver, cardiovascular disorders such as cardiac arrest,  
CC cerebrovascular disorders such as cerebral ischaemia, nervous system  
CC disorders such as Alzheimer's disease, infections caused by bacteria,  
CC viruses and fungi, ocular disorders such as corneal infection, endocrine  
CC disorders such as premature labour and infertility, gastrointestinal  
CC disorders such as Crohn's disease, renal disorders such as  
CC glomerulonephritis and respiratory disorders such as asthma. The  
CC polypeptides can also be used to aid wound healing, to prevent skin aging  
CC due to sunburn, to maintain organs before transplantation, to regenerate  
CC tissues and in chemotaxis. The polypeptides can also be used as a food  
CC additive or preservative to increase or decrease storage capabilities.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.

Query Match 76.5%; Score 632; DB 22; Length 176;  
Best Local Similarity 94.0%; Pred. No. 8.3e-66;  
Matches 126; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 AQIYSVAIFLGINLAAFIIIVFSGSMFYSHQSAITATEIRNOVKKEMILAKRFFFIYF 60  
Db 1 AQIYSVAIFLGINLAAFIIIVFSGSMFYSHQSAITATEIRNOVKKEMILAKRFFFIYF 60

QY 61 TDALCWPIPIFVAKPLSLQVIEPCTITSWVICYSAINSLNPILYTLTRPKEMIERF 120  
Db 61 TDALCWPIPIFVAKPLSLQVIEPCTITSWVICYSAINSLNPILYTLTRPKEMIERF 120

QY 121 WHNYRQKSMDSKG 134  
Db 121 WHNYRQKSMDSKG 134

Search completed: September 25, 2003, 14:59:36  
Job time : 42 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 25, 2003, 13:06:05 : Search time 1645 Seconds  
(without alignments)  
7003.235 Million cell updates/sec

Title: US-10-049-568-1

Perfect score: 474  
Sequence: 1 gccacattattcaagtgc.....ggagatgccacctgagttaa 474

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 1215238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

#### Database :

EST.\*

- 1: em\_estba.\*
- 2: em\_estbam.\*
- 3: em\_estin.\*
- 4: em\_estnu.\*
- 5: em\_estov.\*
- 6: em\_estpl.\*
- 7: em\_estro.\*
- 8: em\_hic.\*
- 9: gb\_esti.\*
- 10: gb\_est2.\*
- 11: gb\_hic.\*
- 12: gb\_est3.\*
- 13: gb\_est4.\*
- 14: gb\_est5.\*
- 15: em\_estfun.\*
- 16: em\_estom.\*
- 17: em\_gss\_hum.\*
- 18: em\_gss\_inv.\*
- 19: em\_gss\_pin.\*
- 20: em\_gss\_vit.\*
- 21: em\_gss\_fun.\*
- 22: em\_gss\_mam.\*
- 23: em\_gss\_mus.\*
- 24: em\_gss\_pro.\*
- 25: em\_gss\_rod.\*
- 26: em\_gss\_ptg.\*
- 27: em\_gss\_vrl.\*
- 28: gb\_gssl.\*
- 29: gb\_gss2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	320.4	67.6	841	13 B0228832	B0228832 AGENCOURT
2	284	59.9	873	13 B0274328	B0274328 603532573
3	236	49.8	686	13 B0284397	B0284397 603866563
4	199.8	42.2	489	13 BX089223	BX089223 BX089223

	5	199.8	42.2	799	29	BZ605862
	6	160.6	33.9	778	13	B0279098
	7	146	30.8	648	14	CB216867
	8	136.4	28.8	744	10	EG304121
C	9	131.4	27.7	758	13	EX084029
	10	130	27.4	631	9	AL955885
	11	98.4	20.8	235	9	AW436170
	12	90	19.0	198	14	CA997681
	13	87.6	18.5	197	10	BF942735
	14	87.6	18.5	197	12	BM280297
	15	71.8	15.1	1078	29	CC244249
C	16	69.4	14.6	286	9	AI375172
	17	63.2	13.3	835	28	BZ184017
	18	60.8	12.8	589	13	B0245340
	19	50.6	10.7	456	10	BF021857
	20	49.6	10.5	574	10	BB632910
	21	49.6	10.5	1307	11	AK039086
	22	44	9.3	450	14	H17265
	23	44	9.3	511	2	HS0095693
	24	43.4	9.2	648	13	BM089852
	25	43.4	9.2	660	13	BW239928
	26	43.4	9.2	680	28	BZ022413
	27	43.4	9.2	726	13	BM090495
C	28	43.4	9.2	733	13	BM086693
C	29	43.4	9.2	747	13	BM090486
	30	43	9.1	602	29	AG228879
C	31	42.6	9.0	572	28	AQ697633
	32	42.4	8.9	1101	29	CNS0169X
C	33	42.2	8.9	896	29	CNS07DBB
	34	42.2	8.9	1201	13	EX446356
	35	41.8	8.8	574	29	BZ387557
C	36	41.8	8.8	696	13	BM160624
	37	41.8	8.8	739	29	BZ390866
	38	41.4	8.7	1101	29	CNS0145U
C	39	41.2	8.7	779	29	EX229966
	40	41	8.6	1101	29	CNS0039G
C	41	40.6	8.6	208	14	CD429485
	42	40.6	8.6	1042	13	EX375647
	43	40.2	8.5	712	13	BM043475
C	44	40.2	8.5	789	13	BM015323
	45	40	8.4	553	28	AQ775843
						HS_2151_B

#### ALIGNMENTS

RESULT 1  
BQ228832  
LOCUS  
DEFINITION  
BQ228832  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

BQ228832  
AGENCOURT\_7522585 NIH\_MGC\_72 Homo sapiens cDNA clone IMAGE:6051251  
5', mRNA sequence.  
BQ228832  
BQ228832.1 GI:20410232  
EST.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 841)  
NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
Tissue Procurement: ATCC/DCTD/DTp  
cDNA Library Preparation: Life Technologies, Inc.  
DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)  
Clone distribution: Agencourt Bioscience Corporation  
found through the I.M.A.G.E. Consortium information can be  
<http://image.llnl.gov>  
Plate: LLAM13304 row: m column: 12  
High quality sequence start: 87

High quality sequence stop: 603.  
 FEATURES  
 source  
 1. 841  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6051251"  
 /tissue\_type="melanotic melanoma"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NH.MGC.72"  
 /note="Organ: skin; Vector: pCMV-SPORT6; Site.1: NotI;  
 Site.2: SalI; Cloned unidirectionally. Primer: Oligo dt.  
 Average insert size 2 kb. Library constructed by Life  
 Technologies."  
 BASE COUNT 240 a 160 c 170 g 269 t 2 others  
 ORIGIN

Query Match 67.6%; Score 320.4; DB 13; Length 841;  
 Best Local Similarity 95.4%; Pred. No. 7.9e-62;  
 Matches 330; Conservative 0; Mismatches 16; Indels 0; Gaps 0;  
 QY 1 GCCAGATTATTCAGTGGCAATTTTCCTGTTATTAATTTGGCGCATTTATCATCATATA 60  
 DB |||||  
 QY 305 GCCAGATTATTCAGTGGCAATTTTCCTGTTATTAATTTGGCGCATTTATCATCATATA 364  
 DB |||||  
 QY 61 GTTTTTCCTATGGAAGCATGTTTATAGTGTTCATCAAGTGCATTAACAGCAACTGAA 120  
 DB |||||  
 QY 365 GTTTTTCCTATGGAAGCATGTTTATAGTGTTCATCAAGTGCATTAACAGCAACTGAA 424  
 DB |||||  
 QY 121 ATACGGAATCAAGTTAAAGAGATGATCCTTGGCAACGTTTTCCTTTATAGTATTT 180  
 DB |||||  
 QY 425 ATACGGAATCAAGTTAAAGAGATGATCCTTGGCAACGTTTTCCTTTATAGTATTT 484  
 DB |||||  
 QY 181 ACTGATGCAATTATGCTGATACCATTTTGTAGCGAAACCTCTTTCACATGCTTCAGGTA 240  
 DB |||||  
 QY 485 ACTGATGCAATTATGCTGATACCATTTTGTAGCGAAACCTCTTTCACATGCTTCAGGTA 544  
 DB |||||  
 QY 241 GAATATCAGAGTACCATACCTCTGGGTAGTGGTATTTCTGCGCATTAACAGTCT 300  
 DB |||||  
 QY 545 GAATATCAGAGTACCATACCTCTGGGTAGTGGTATTTCTGCGCATTAACAGTCT 604  
 DB |||||  
 QY 301 TTGAACCCCAATTCCTATACCTCTGACCACCAAGACCATTTAAAGAA 346  
 DB |||||  
 QY 605 TTGAACCCCAATTCCTATACCTCTGACCACCAAGACCATTTAAAGAA 650  
 DB |||||

RESULT 2  
 BU274328 873 bp mRNA linear EST 26-NOV-2002  
 LOCUS 603532573F1 CSEQCHN53 Gallus gallus CDNA clone CHEST489a19 5', mRNA  
 DEFINITION  
 ACCESSION  
 VERSION BU274328.1 GI:25545278  
 KEYWORDS EST.  
 SOURCE Gallus gallus (chicken)  
 ORGANISM Gallus gallus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
 Phasianinae; Gallus.  
 1 (bases 1 to 873)  
 Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,  
 Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.  
 A Comprehensive Collection of Chicken cDNAs  
 Curr. Biol. 12 (22), 1965-1969 (2002)  
 22335534  
 PubMed 12445392  
 Contact: Simon Hubbard  
 Department of Biomolecular Sciences  
 University of Manchester Institute of Science and Technology (UMIST  
 )  
 PO Box 88, Manchester, M60 1QD, UK  
 Tel: 01612008930  
 Fax: 01612360409

Email: Simon.Hubbard@umist.ac.uk.  
 FEATURES  
 Location/Qualifiers  
 1..873  
 /organism="Gallus gallus"  
 /mol\_type="mRNA"  
 /strain="Compton Line 151"  
 /db\_xref="taxon:9031"  
 /clone="ChESP489a19"  
 /sex="Female"  
 /tissue\_type="cerebrum"  
 /dev\_stage="adult"  
 /lab\_host="DH10B"  
 /clone\_lib="CSEQCHN53"  
 /note="Organ: brain; Vector: pBluescript II KS(+); Site.1:  
 EcoRI; Site.2: NotI; This normalized library was  
 constructed from 1 million independent clones. CDNA  
 synthesis was initiated using an oligo(df) primer, using  
 methylated C in the first strand synthesis reaction.  
 Following this first strand reaction, double-stranded cDNA  
 was blunted, ligated to NotI adapters, digested with EcoRI  
 , size-selected, and cloned into the NotI and EcoRI  
 compatible sites of a custom modified MCS of the  
 pBluescript (KS+) vector. The library was normalized in 2  
 rounds using conditions adapted from Soares et al., PNAS  
 (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6  
 (1996): 791, except that a significantly longer  
 reannealing hybridization was used."  
 BASE COUNT 252 a 190 c 160 g 271 t  
 ORIGIN

Query Match 59.9%; Score 284; DB 13; Length 873;  
 Best Local Similarity 79.2%; Pred. No. 1.2e-53;  
 Matches 374; Conservative 0; Mismatches 95; Indels 3; Gaps 3;  
 QY 4 CAGATTTATTCAGTGGCAATTTTCCTGTTATTAATTTGGCGCATTTATCATCATAGTT 63  
 DB |||||  
 QY 42 CAGATTTATTCAGTGGCAATTTTCCTGTTATTAATTTGGCGCATTTATCATCATAGTT 99  
 DB |||||  
 QY 64 TTTTCTATGGAAGCATGTTTATAGTGTTCATCAAAAGTCCCATTAACAGCAACTGAAATA 123  
 DB |||||  
 QY 100 TTTTCTATGGAAGCATGTTTATAGTGTTCATCAAAAGTCCCATTAACAGCAACTGAAATA 159  
 DB |||||  
 QY 124 CGAATCAAGTTAAAGAGATGATCCTTGGCAACGTTTTCCTTTATAGTATTTACT 183  
 DB |||||  
 QY 160 CAGAATCATATTTAAAGAGATGATCCTTGGCAACGTTTTCCTTTATAGTATTTACT 219  
 DB |||||  
 QY 184 GATGATTTATGCTGGATACCATTTTGTAGCAAAACCTCTTTCACTGCTTCAGTAGAA 243  
 DB |||||  
 QY 220 GATGATTTATGCTGGATACCATTTTGTAGCAAAACCTCTTTCACTGCTTCAGTAGAA 279  
 DB |||||  
 QY 244 ATACAGTACCATTAACCTCTGGGTAGTGGTATTTCTGCCATTAACAGTGTCTTG 303  
 DB |||||  
 QY 280 ATACAGTACCATTAACCTCTGGGTAGTGGTATTTCTGCCATTAACAGTGTCTTG 339  
 DB |||||  
 QY 304 AACCAATTTCTATCTACTCTGACCAAGACCATTTAAAGAAATGATTCATCGTTTGG 363  
 DB |||||  
 QY 340 AATCTCTCTCTACATTTGTACTACAGACCATTTCAAGAAATGATTCACCATTTGG 399  
 DB |||||  
 QY 364 CATACTACAGCAAAAGAAATCTATGGACAGCAAGGTA-TCAGAAACATATAGTCTCA 422  
 DB |||||  
 QY 400 TACAACATACAGCAAAAGAAATCTATGGACAGCAAGGTA-TCAGAAACATATAGTCTCA 459  
 DB |||||  
 QY 423 TCATTTCATCTGGGGGAATGTGGCCACTGCAGGAGATGCCACCTGAGTTAA 474  
 DB |||||  
 QY 460 TCCTTCATTTGGGTAGAGATGTGGCCAAATGCATGAAATCACACCAAGCTAA 511  
 DB |||||

RESULT 3  
 BU284397 686 bp mRNA linear EST 27-NOV-2002  
 LOCUS 603866563F1 CSEQCHN54 Gallus gallus CDNA clone ChEST886m16 5', mRNA  
 DEFINITION  
 ACCESSION  
 BU284397

```

VERSION      BU284397.1  GI:25733853
KEYWORDS     EST.
SOURCE       Gallus gallus
ORGANISM     Gallus gallus
              Archosauaria; Aves; Neognathae; Galliformes; Phasianidae;
              Phasianinae; Gallus.
REFERENCE    1 (bases 1 to 686)
AUTHORS      Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E.,
              Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.
              A Comprehensive Collection of Chicken cDNAs
              Curr. Biol. 12 (22), 1965-1969 (2002)
JOURNAL      22335534
MEDLINE      12445392
PUBMED
COMMENT      Contact: Simon Hubbard
              Department of Biomolecular Sciences
              University of Manchester Institute of Science and Technology (UMIST)
              )
              PO Box 88, Manchester, M60 1QD, UK
              Tel: 01612008930
              Fax: 01612360409
              Email: Simon.Hubbard@umist.ac.uk.
              Location/Qualifiers
              1. 686
              /organism="Gallus gallus"
              /mol_type="mRNA"
              /strain="Compton Line 151"
              /db_xref="taxon:9031"
              /clone="CHS7886m16"
              /sex="Female"
              /tissue_type="not cerebrum or cerebellum"
              /dev_stage="adult"
              /lab_host="DH10B"
              /clone_lib="CSECHN54"
              /note="Organ: brain; Vector: pBluescript II KS(+); Site:1:
              EcoRI; Site:2: NotI; This normalized library was
              constructed from 1 million independent clones. cDNA
              synthesis was initiated using an oligo(dT) primer, using
              methylated C in the first strand synthesis reaction.
              Following this first strand reaction, double-stranded cDNA
              was blunt-ended, ligated to NotI adapters, digested with EcoRI
              , size-selected, and cloned into the NotI and EcoRI
              compatible sites of a custom modified MCS of the
              pBluescript (KS+) vector. The library was normalized in 2
              rounds using conditions adapted from Soares et al., PNAS
              (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
              (1996): 791, except that a significantly longer
              reannealing hybridization was used."
BASE COUNT   207 a 151 c 131 g 196 t
ORIGIN
Query Match 49.8%; Score 236; DB 13; Length 686;
Best Local Similarity 78.4%; Pred.No. 7.8e-43;
Matches 345; Conservative 0; Mismatches 90; Indels 5; Gaps 5;

QY 36 TAATTGGCGCATTTATCATCATAGTTTTTTTCTTATGGAAGCATGTTTTATAGTTGCA 95
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2 TAAACTGGCAGCTGTCTCATCAT- GCTTTTCTATGGAAGTATGTTCTACAGTTTCA 60

QY 96 TCAAGTGGCATACAGCACTGAATACGAATCAGTTTAAAAAGAGATGATCTTGC 155
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 61 CCAACAGCTATTATGGCCACTGAATTCAGATCATATATAAAGAGATGATAC-TGC 119

QY 156 CAACAGTTTTTCTTTATATATTTTACTGATTCATTATGCTGGATACCCATTTTGTAGC 215
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 120 CAACAGCTTTTCTTCATGTTGATTTACCGATGCACTTTG-TGGATACCAATTTTATTT 178

QY 215 GAAACCTCTTCTACCTGCTTCAGGATGAGAAATPACAGATACCAATCTTTGGTAGTGAT 275
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 179 GAAACCTCTTCTTCTTACTGCAAGTAGAATPACAGATACCAATCATC-TGGGTAGTGAT 237

QY 276 TGGTTATTCWGCATTAACAGTGTCTTGAACCAATTCCTATACTCTGACCAAGACC 335
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

```

Db 238 TTTTATCTTGCCAAATAAATAGTGTCTCTGAATCCCTCTCTACACTTTGACTACACGACC 297
QY 336 ATTAAAGAATGATTCATCGGTTTGGCATTACTACAGACAAAGAAATATATGGACAG 395
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 298 CTTCAAGAATGATTCACCAGTTTGTACACTACAGACAAAGATCCAAGAGAGG 357
QY 396 CAAAGTA-TCAGAAACATATGCTCCATCATCTGGGGGGAATGTGCCACTGCA 454
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 358 CAAAGGCAGTCAGAAAGCTTATGTTCCATCTTCTTATTTGGTAGATGTGCCAATGCA 417
QY 455 GGAGATGCCACCTGAGTTAA 474
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 418 TGAATTCACCAAGCTAA 437

RESULT 4
LOCUS      EX089223          489 bp      mRNA      linear      EST 23-JAN-2003
DEFINITION EX089223 Soares_NHMPu_S1 Homo sapiens cDNA clone IMAGE998M245082 ;
IMAGE:2063423, mRNA sequence.
ACCESSION  EX089223
VERSION     EX089223.1  GI:27825935
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 489)
AUTHORS     Ebert, L., Heil, O., Hennig, S., Neubert, P., Partsch, E., Peters, M.,
              Radelof, U., Schneider, D. and Korn, B.
              Human Unigeneset - RZPD3
              Unpublished
              Contact: Ina Rolfs
              RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
              Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
              RZPD; IMAGE998M245082.
              RZPDLIB; I.M.A.G.E. cDNA Clone Collection;
              Human Unigeneset - RZPD3 (RZPDLIB No.972)
              http://www.rzpd.de/CloneCards/cgi-
              bin/showLib.pl.cgi/response?libNo=972 Contact: Ina Rolfs
              RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
              Heubnerweg 6, D-14059 Berlin, Germany
              Tel: +49 30 32639 101
              Fax: +49 30 32639 111
              www.rzpd.de
              This clone is available royalty-free from RZPD:
              contact RZPD (clone@rzpd.de) for further information. Seq primer:
              M13r, Primer sequence: TTTCACAGGAACACGATGAC.
FEATURES    Location/Qualifiers
             1..489
             /organism="Homo sapiens"
             /mol_type="mRNA"
             /db_xref="taxon:9606"
             /clone="IMAGE998M245082 ; IMAGE:2063423"
             /tissue_type="pooled human melanocyte, fetal heart, and
             pregnant uterus"
             /lab_host="DH10B"
             /clone_lib="Soares_NHMPu_S1"
             /note="Organ: mixed (see below); Vector: pT73D-Pac
             (Pharmacia) with a modified polylinker; Site.1: Not I;
             Site.2: Eco RI; Equal amounts of plasmid DNA from three
             normalized libraries (melanocyte 2NHM, pregnant uterus
             NDHPU, and fetal heart NDH19W) were mixed, and ss circles
             were made in vitro. Following HAP purification, this DNA
             was used as tracer in a subtractive hybridization
             reaction. The driver was PCR-amplified cDNAs from pools of
             5,000 clones made from the same 3 libraries. The pools
             consisted of I.M.A.G.E. clones 260232-265223,
             340488-345479, and 484488-489479."
BASE COUNT  166 a 95 c 86 g 142 t
ORIGIN
Query Match 42.2%; Score 199.8; DB 13; Length 489;

```

```

Best Local Similarity 88.7%; Pred. No. 1.1e-34;
Matches 228; Conservative 0; Mismatches 27; Indels 2; Gaps 1;

QY 218 AACCTCTTTCAGTCTGTTGAGCCAAATACAGTACCAATCTCTGGTAGTAGTG 277
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 147 AAACACTAAACATATTCTATTTCTTTTGTAGGTACCAATCTCTGGTAGTAGTT 206
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 278 GTATTCCTGCCATTAAACAGTCTTTGAAACCAATCTCTATATCTCTGACCAAGACCAT 337
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 207 TTATTCCTGCCATTAAACAGTCTTTGAAACCAATCTCTATATCTCTGACCAAGACCAT 266
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 338 TTAAGAAGATGATTCATCGGTTTGGCATAACTACAGACAAAGAAATCTATGACAGCA 397
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 267 TTAAGAAGATGATTCATCGGTTTGGCATAACTACAGACAAAGAAATCTATGACAGCA 326
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 398 AAGTATACAGAAACATATGTCATCATCTCTGGGGGGAATGTGGCCATGACAGCA 457
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 327 AAGG--TCAGAAACATATGTCATCATCTCTGGTGGAAATGTGGCCACTGCAGCA 384
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 458 GATGCCACCTGAGTTAA 474
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 385 GATGCCACCTGAGTTAA 401
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 5
BZ605862
LOCUS WHACX47TF Human MCF7 breast cancer cell line library (MCF7_1) Homo
DEFINITION sapiens genomic clone MCF7_1-18H22, genomic survey sequence.
ACCESSION BZ605862
VERSION BZ605862.1 GI:31514423
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 799)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
TITLE Volik S., Zhao S., Chin K., Brebner, J.H., Herndon, D.R., Tao, Q.,
JOURNAL Kowbel, D., Huang, G., Lapuk, A., Kuo, W.-L., Magrane, G., de Jong, P.,
MEDLINE Gray, J.W. and Collins, C.
PUBMED ESP: a sequence-based approach to the structural genomics of tumors
12445392
COMMENT Unpublished (2002)
CONTACT: Volik SV
JOURNAL Colin Collins' lab
COMMENT UCSF Comprehensive Cancer Center
UCSF Box 0808, San Francisco, CA 94143-0808, USA
Tel: 415 502 7066
Fax: 415 502 5665
Email: svolik@cc.ucsf.edu
This clone is available from Amplicon Express
http://www.genomex.com
Class: BAC ends.
FEATURES
source
Location/Qualifiers
1..799
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="MCF7_1-18H22"
/sex="female"
/clone_lib="Human MCF7 breast cancer cell line library
(MCF7_1)"
/notes="Vector: pECBAC1; Site_1: HindIII; This library was
constructed from MCF7 breast cancer cell line by Amplicon
Express (http://www.genomex.com) using their standard
procedure."
BASE COUNT 277 a 145 c 127 g 250 t
ORIGIN
Query Match 42.2%; Score 199.8; DB 29; Length 799;
Best Local Similarity 86.7%; Pred. No. 1.1e-34;
Matches 228; Conservative 0; Mismatches 27; Indels 2; Gaps 1;

QY 218 AACCTCTTTCAGTCTGTTGAGCCAAATACAGTACCAATCTCTGGTAGTAGTG 277

```

```

||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 389 AAACACTAAACATATTCTATTTCTTTTGTAGGTACCAATCTCTGGTAGTAGTTT 448
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 278 GTATTCCTGCCATTAAACAGTCTTTGAAACCAATCTCTATATCTCTGACCAAGACCAT 337
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 449 TTATTCCTGCCATTAAACAGTCTTTGAAACCAATCTCTATATCTCTGACCAAGACCAT 508
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 338 TTAAGAAGATGATTCATCGGTTTGGCATAACTACAGACAAAGAAATCTATGACAGCA 397
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 509 TTAAGAAGATGATTCATCGGTTTGGCATAACTACAGACAAAGAAATCTATGACAGCA 568
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 398 AAGTATACAGAAACATATGTCATCATCTCTGGGGGGAATGTGGCCATGACAGCA 457
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 569 AAGG--TCAGAAACATATGTCATCATCTCTGGTGGAAATGTGGCCACTGCAGCA 626
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 458 GATGCCACCTGAGTTAA 474
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 627 GATGCCACCTGAGTTAA 643
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 6
BZ279098
LOCUS 603865023F1 CSEQCHN54 Gallus gallus cDNA clone CHEST82n6 5', mRNA
DEFINITION sequence.
ACCESSION BZ279098
VERSION BZ279098.1 GI:25728552
KEYWORDS EST.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
REFERENCE 1 (bases 1 to 778)
AUTHORS Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E.,
Fong, W.T., Fickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.
TITLE A Comprehensive Collection of Chicken cDNAs
JOURNAL Curr. Biol. 12 (22), 1965-1969 (2002)
MEDLINE 22335534
PUBMED 12445392
COMMENT Contact: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology (UMIST)
)
PO Box 88, Manchester, M60 1QD, UK
Tel: 01612008930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.
Location/Qualifiers
1..778
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="Compton Line 151"
/db_xref="taxon:9031"
/clone="CHEST82n6"
/sex="Female"
/tissue_type="not cerebrum or cerebellum"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="CSEQCHN54"
/notes="Organ: brain; Vector: pBluescript II KS(+); Site_1:
EcoRI; Site_2: NotI; This normalized library was
constructed from 1 million independent clones. cDNA
synthesis was initiated using an oligo(dT) primer, using
methyated c in the first strand synthesis reaction.
Following this first strand reaction, double-stranded cDNA
was blunted, ligated to NotI adapters, digested with EcoRI
, size-selected, and cloned into the NotI and EcoRI
compatible sites of a custom modified MCS of the
pBluescript (KS+) vector. The library was normalized in 2
rounds using conditions adapted from Soares et al., PNAS
(1994) 91: 9228-9232 and Donaldo et al., Genome Research 6
(1996): 791, except that a significantly longer

```

1  
 2  
 3  
 4  
 5  
 6  
 7  
 8  
 9  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525

	pyresen	nitrogen	corpo
BASE COUNT	180 a	112 c	223 t
ORIGIN		133 g	

```
Query Match      30.8%; Score 146; DB 14; Length 648;
Best Local Similarity 100.0%; Pred.No. 1.4e-22;
Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy 1 GCCCAGATTTATTCAGTGGCAATTTTCTTGGTATTAATTTGGCCGCATTTATCATCATCA 60

D**b** 503 GCCCAGATTATTTCAGTGGCAATTTTCTTTGGTATTAAATTGGCCGATTTATCATATA 562

QY 61 GTTTTTCCTATGGAAGCATGTTTATTAGTGTTCAATCAAAGTGCATACAGCAACTGAA 120

Db 563 GTTTTTCCTATGGAAGCATGTTTTATAGTGTTCATCAAAGTCCATAACAGCAACTGAA 622

QY 121 ATACGGAATCAAGTTAAAAAAGAGAT 146

D<sub>b</sub> 623 ATACGGAATCAAGTTAAAAAAGAGAT 648

BG304121/c

**DEFINITION** fl30f05.x1 Sugano Kawakami zebrafish DRA Danio rerio cDNA clone IMAGE:2815552 2' similar to mp.000000.000000 CC5042 PROTEIN.

ACCESSION BC304121 mRNA sequence.

VERSION	EG304121.1	GI:13101648
KEYWORDS	EST.	
SOURCE	Danio rerio (zebrafish)	
ORGANISM	Danio rerio	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariocphys; Cypriniformes; Cyprinodontiformes	

REFERENCE  
1 (bases 1 to 744)  
CYPRINIDAE; DANIO.

**AUTHORS**  
Sugano, S., Kawakami, K., Johnson, S., Li, F., Maria, M., Eddy, S.,  
Hillier, L., Clifton, S., Allen, M., Gibbons, M., Jost, S., Kucaba, T.,

Marth, J., Pape, D., Steptoe, M., Underwood, K., Theising, B., Ritter, E., Bowers, Y., Wylie, T., Waterston, R. and Wilson, R.

ILLI  
JOURNAL  
WASNO ZEDRATISH EST PROJECL 1999  
Unpublished

COMMENT  
Contact: S.E. JOHNSON  
Washington University School of Medicine

4444 FOREST PARK PARKWAY, BOX 8301, ST. LOUIS, MO 63106, USA  
Tel: 314 286 1800

FAX: 314 200 1010  
Email: [est@watson.wustl.edu](mailto:est@watson.wustl.edu)

Sequencing by: Washington University Genome Sequencing Center Clonally  
Library constructed by Dr. Sumio Sugano and Dr. Koichi Kawarada at  
University of Tokyo

Consortium/LLNL, send email to: [info@image.llnl.gov](mailto:info@image.llnl.gov)

High quality sequence stop: 510.

source 1.744

```

/mol_type="mRNA"
/organism="Drosophila"

```

```
/db_xref="taxon:7955"
```

```
/sex="mixed (one male and one female, including
```

```

/dev_stage="adult"

```

```
/clone_lib="Sugano Kawakami zebrafish DRA"
```

Site\_2: DraIII (CACCATGTG); 1st strand cDNA was primed

double-stranded cDNA was ligated to a DraIII adaptor

sites of the pME18S-FL3 vector (5' site CACTGTGTG, 3' site

INVESTIGATION OF THE EFFECTS OF THE 1960-1961

insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science) and kindly donated by Dr. Koichi Kawakami. Custom primers for sequencing: 5' end primer CTTCGTCCTAAAGCTGCG and 3' end primer GCACTCGAGCTGACACA.

```

BASE COUNT      204 a 158 c 186 g 196 t
ORIGIN

Query Match      28.8%; Score 136.4; DB 10; Length 744;
Best Local Similarity 62.1%; Pred. No. 2e-20;
Matches 215; Conservative 0; Mismatches 131; Indels 0; Gaps 0;

QY 25 TTCTTGTTGTTAAATTTGGCGCATTTATCATCATAGTTTTCCTATGGAACATCTTT 84
   || || || || || || || || || || || || || || || || || || || ||
Db 744 TTCATGGGACTTACCGGGTGCTTCTCATCATCGTGGTCTGCTACTCCAGCANGTTT 685

QY 85 TATAGTGTTCATCAAGTGCATACAGACACTGAATACGGAATCAAGTTAAAAAGAG 144
   || || || || || || || || || || || || || || || || || || || ||
Db 684 TATCCATCTATAAGACGGGCATCAGAGCCACAGAGCTGCGGGTGCAGTGCACAGAGAC 625

QY 145 ATGATCCTGGCCAAACGTTTTTCTTTATAGTATTTACGTAGTCATTATGCTGATACCC 204
   || || || || || || || || || || || || || || || || || || || ||
Db 624 GTGCCATGGCACACCGTTCTTCTCATCTGTTTCTGATGCCCTCTGCTGATCCCC 565

QY 205 ATTTTGTAGCGAAACCTTTTCACTGCTTTCAGGTAGAAATACCAAGGTACCATACTCT 264
   || || || || || || || || || || || || || || || || || || || ||
Db 564 ATATTCATGTCATAAATCCTCTCTAATGGAGTGGAGATACCAAGGAACCATCATCT 505

QY 265 TGGGTAGTATGGTTATCTGCCATTAACAGTGCCTTTGACCCAAATCTCTATATCTG 324
   || || || || || || || || || || || || || || || || || || || ||
Db 504 TGGGTGATGTTTCACTCCTGCCATCAACAGTGCCTTTAAACCCATCTTTACACACTG 445

QY 325 ACCAAGACACATTTAAAGAAATGATTCATCGTGGTTTGGCATACT 370
   || || || || || || || || || || || || || || || || || || || ||
Db 444 ACCAAGATTTTTCAGAGACGAGTGGACATTTATGGGCACT 399

```

```

RESULT 9
BX084029
LOCUS
DEFINITION
BX084029 AGENAE Rainbow trout normalized multi-tissues library
(tcac) Oncorhynchus mykiss cDNA clone tcac0004c.h.01 5prim, mRNA
sequence.
BX084029
BX084029.1 GI:27747446
EST.
Oncorhynchus mykiss (rainbow trout)
Oncorhynchus mykiss

```

```

Oncorhynchus mykiss (rainbow trout)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei;
Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
1 (bases 1 to 758)
Govoroun,M., Guiguen,Y. and Le Gac,F.
Construction and primary characterization of normalized cDNA
libraries in rainbow trout, Oncorhynchus mykiss
Unpublished
Contact: Guiguen Y
INRA - SCRIBE
Campus de Beaulieu, RENNES cedex, 35042, France
Tel: 02.23.48.50.09
Fax: 02.23.48.50.20
Email: Yann.Guiguen@beaulieu.rennes.inra.fr
Sequence cleaned of vector, adaptor and repetitions. Contact us
at signasupport@jouy.inra.fr to obtain the chromatogram of this
sequence.
Seq primer: T7
Location/Qualifiers
1. .758
/organism="Oncorhynchus mykiss"
/moi_type="mRNA"
/db_xref="taxon:8022"

```

```

FEATURES
source
1. .631

```

```

/clone="tcac0004c.h.01"
/tissue_type="adipose tissue, blood, brain,
differentiating gonads, gills, interrenal, intestine,
kidney, liver, muscle, ovary, pituitary, testis"
/dev_stage="from embryos to adults"
/lab_host="DH10B"
/clone_lib="AGENAE Rainbow trout normalized multi-tissues
library (tcac)"
/notes="Vector: pT73D-pac; Clone distribution : AGENAE
Resource centre. Francois Pluimi,
Francois.Plumi@jouy.inra.fr, INRA, CERA Radiobiologie et
Etude du genome (LREG), Domaine de Vilvert, 78352,
Jouy-en-Josas cedex, France"
BASE COUNT      165 a 217 c 178 g 198 t
ORIGIN

```

```

Query Match      27.7%; Score 131.4; DB 13; Length 758;
Best Local Similarity 61.6%; Pred. No. 2.7e-19;
Matches 210; Conservative 0; Mismatches 131; Indels 0; Gaps 0;

QY 10 TATTCAGTGGCAATTTTCTTGGTATTAAATTTGGCGCATTTATCATCATAGTTTTCCTC 69
   || || || || || || || || || || || || || || || || || || || ||
Db 282 TATCTATCGGATATTCCTGGGTCTGACCTGGCAGATTCTCTAGTATCGTGTCTCT 341

QY 70 TATGGAAGCATCTTTTATAGTGTTCATCAAAAGTGCCATACACGCAACTGAAATACGGAAT 129
   || || || || || || || || || || || || || || || || || || || ||
Db 342 TACTCTCCATGTCTTACTCCATCTACAAGACGTGGTATCAACGCTACAGACCTGAGGAGC 401

QY 130 CAAGTTAAAAAGAGATGATCCTTGGCCAAACGTTTTTCTTTATAGTATTTATCTGATGCA 189
   || || || || || || || || || || || || || || || || || || || ||
Db 402 AGGCTACACAGAGACGTGGCTGAGCTTAACAGGTCTTCTTCATCGTCTTCTCTGACGC 461

QY 190 TTATGCTGGATCCCATTTTGTAGCGAAACCTTTTCACTGCTTCAGGTAGAAATACCA 249
   || || || || || || || || || || || || || || || || || || || ||
Db 462 CTCTGCTGGATCCCATCTTCTGCTGACAGATCTGCTGCTGACAGTATACCA 521

QY 250 GGTACCATACCTCTTTGGGTAGTGAATGTTATCTGCAATTAACAGTGTGTTTGAACCCA 309
   || || || || || || || || || || || || || || || || || || || ||
Db 522 GGTACCATCTCATCTCTGGTGGTGTATCTTTATATACCATCAACAGTGCCTGAACCCC 581

QY 310 ATTCTCTATACCTCACCACAGACCATTTTAAAGAAATGAT 350
   || || || || || || || || || || || || || || || || || || || ||
Db 582 ATCTCTACCTCTGACTACCAGCTTCTTCAGAGAACAGGT 622

```

```

RESULT 10
AL955885
LOCUS
DEFINITION
AL955885 XSC-gastrula Silurana tropicalis cDNA clone TCas123K02 5',
mRNA sequence.
AL955885
AL955885.1 GI:25779480
ACCESSION
VERSION
KEYWORDS
SOURCE
EST.
ORGANISM
Silurana tropicalis (western clawed frog)
Silurana tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
Xenopodinae; Silurana.

```

```

REFERENCE
1 (bases 1 to 631)
Taylor,R., Ashurst,J.L., Croning,M.D.R., Zorn,A.M. and Rogers,J.
Sanger Xenopus tropicalis EST project 2002
Unpublished
Contact: Taylor R
Sanger Centre
Hinxtion, Cambridgeshire, CB10 1SA, UK
Email: trop@sanger.ac.uk
Sanger Xenopus tropicalis EST project 2001
TROPICALIS_SEQUENCE_ID: TCas123K02.picSP6
Sequencing primer: SP6
This sequence is from a Xenopus Gene Collection (XGC) library
constructed by Aaron M. Zorn.
Location/Qualifiers
1. .631

```

```

FEATURES
source
1. .631

```

```

; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 1232
; LENGTH: 1557
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (16)
; OTHER INFORMATION: Identity of nucleotide at the above locations are unknown.
US-09-134-001C-1232

Query Match      8.0%; Score 37.8; DB 4; Length 1557;
Best Local Similarity 54.7%; Pred. No. 0.5;
Matches 75; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY 337 TTTAAGAAATGATTCATCGGTTTGGCATCTACTACACACAAAGAAATCTATGGACAGC 396
Db 1293 TTTAAGAAATGATTCATCGGTTTGGCATCTACTACACACAAAGAAATCTATGGACAGC 396

QY 397 AAAGGTATCAGAAACATATGCTCCATCATCTGGGGGAAATGTGGCCACATGCAGG 456
Db 1233 AATAAGATTAATCTTCAATGTTTCATTAATTTTCTCCAAATGCTAGCTTCAAT 1174

QY 457 AGATGCCACCTGAGTGA 473
Db 1173 ATATGTCGATGAATCA 1157

```

Search completed: September 25, 2003, 13:59:13  
Job time : 59 secs



```

/organism="Sus scrofa"
/mol_type="mRNA"
/db_xref="taxon:9823"
/clone="pIGR8"
/sex="female"
/tissue_type="uterus"
/dev_stage="neonatal day 14"
/clone_lib="Sus scrofa neonatal uterus RT-PCR"
/notes="Organ: uterus; cDNA generated by RT-PCR from total
RNA extracted from porcine neonatal uterine tissue."
BASE COUNT      42 a   60 c   36 g   60 t
ORIGIN
Query Match      19.0%; Score 90; DB 14; Length 198;
Best Local Similarity 66.5%; Pred. No. 5.8e-10;
Matches 129; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

QY 178 TTATCTGATGATGATGCTGGATACCACTTTTGTAGGGAACCTCTTCATCGTTTCAG 237
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2 TTCTCTGATGCACTCTGCTGATCCGCTATTTCATTAATAATCTCCCTCTTCGGG 61
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 238 GTAGAAATACAGGTACCAACCTCTTGGGTAGTGATTGTTATTCGCCATTACAGT 297
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 62 GTGGAATACAGGCTACATCTCTCTGGGTGGTCATTTTTCCTCGGTGACAGT 121
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 298 GCTTTGAACCAATCTCTACTCTGACCAACAGCACTTTAAAGAAATGATTCATCGG 357
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 122 GCTTGAACCAATCTCTACTCTGACCAACAGCACTTTTCAAGGACAGTTGAACAG 181
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 358 TTTTGGCATACTA 371
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 182 CTGCTGCACACCA 195
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 13
BF942735
LOCUS
DEFINITION
EST-CD15N-040 human CD15+ myeloid progenitor cells cDNA library
BF942735
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
REFERENCE
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 197)
AUTHORS
Chen, J., Lee, S., Zhou, G., Rowley, J.D. and Wang, S.M.
TITLE
A high-throughput GLGI procedure for converting large number of
SAGE tag sequences into 3' ESTs
JOURNAL
Unpublished
COMMENT
Contact: Wang SM
Hem/Onc
University of Chicago Medical Center
5841 S. Maryland Ave., MC2115, Chicago, IL 60637, USA
Tel: 773-702-6788
Fax: 773-702-3002
Email: swanglemidway.uchicago.edu
This EST fragment was amplified from human CD15+ myeloid progenitor
cells cDNA library with GLGI technique (Generation of Longer cDNA
fragments from SAGE tags for Gene Identification, Proc. Natl. Acad.
Sci. USA 97, 349, 2000), which starts from the 3' end till the last
CATG site of the target cDNA sequence.
Seq primer: M13 Forward.
Location/Qualifiers
1..197
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="bone marrow"
/cell_type="CD15+ myeloid progenitor cells"
/clone_lib="human CD15+ myeloid progenitor cells cDNA
Library"
FEATURES
source
Location/Qualifiers
1..197
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="bone marrow"
/cell_type="CD15+ myeloid progenitor cells"
/clone_lib="human CD15+ myeloid progenitor cells cDNA
Library"

```

```

BASE COUNT      44 a   38 c   37 g   78 t
ORIGIN
Query Match      18.5%; Score 87.6; DB 10; Length 197;
Best Local Similarity 66.3%; Pred. No. 2e-09;
Matches 126; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 62 TTTTTCCTATGGAAGCATCTTTTATAGTGTTCATCAAGTGCCATAACAGCAACTGAAA 121
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 8 TTTTCCCTATATACATATGTTCTGTTCATCAAAAACCGCTTGCACACACAGAG 67
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 122 TACGAATCAAGTTAAAAAGAGATGATCCTTGGCCAAACCTTTTCTTTATAGTATTA 181
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 68 TAAGGAATTTTGTGGAAGAGAGTGGCTGTGCAAAATCGTTCTTTTATAGTGTCT 127
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 182 CTGATGCATTATGCTGATACCACTTTTGTAGAGGAACCTTTTCACCTCCTCAGGTAG 241
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 128 CTGATGCCATCTGCTGATTCCTGTATTGTAGTAAATCCTTCCCTCTTCGGGTGG 187
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 242 AAATACCAAGG 251
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 188 AATACCATG 197
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 14
BM280297
LOCUS
DEFINITION
3'EST-NCI15-045 human CD15+ myeloid progenitor cells cDNA library
Homo sapiens cDNA 3', mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
REFERENCE
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 197)
AUTHORS
Chen, J., Sun, M., Lee, S., Zhou, G., Rowley, J.D. and Wang, S.M.
TITLE
Identifying novel transcripts and novel genes in the human genome
by using novel SAGE tags
JOURNAL
Proc. Natl. Acad. Sci. U.S.A. 99 (19), 12257-12262 (2002)
COMMENT
Contact: Wang SM
Hem/Onc
University of Chicago Medical Center
5841 S. Maryland Ave., MC2115, Chicago, IL 60637, USA
Tel: 773-702-6788
Fax: 773-702-3002
Email: swanglemidway.uchicago.edu
This EST fragment was amplified from human CD15+ myeloid progenitor
cells cDNA library with GLGI technique (Generation of Longer cDNA
fragments from SAGE tags for Gene Identification, Proc. Natl. Acad.
Sci. USA 97, 349, 2000; A high-throughput GLGI procedure for
converting a large number of SAGE tag sequences into 3' ESTs, Genes
& Chromosomes & Cancers, in press), which starts from the 3' end
till the last CATG site of the target cDNA sequence.
Seq primer: M13 Forward.
Location/Qualifiers
1..197
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="bone marrow"
/cell_type="CD15+ myeloid progenitor cells"
/clone_lib="human CD15+ myeloid progenitor cells cDNA
Library"
FEATURES
source
Location/Qualifiers
1..197
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="bone marrow"
/cell_type="CD15+ myeloid progenitor cells"
/clone_lib="human CD15+ myeloid progenitor cells cDNA
Library"
BASE COUNT      44 a   38 c   37 g   78 t
ORIGIN
Query Match      18.5%; Score 87.6; DB 12; Length 197;
Best Local Similarity 66.3%; Pred. No. 2e-09;
Matches 126; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 62 TTTTTCCTATGGAAGCATCTTTTATAGTGTTCATCAAGTGCCATAACAGCAACTGAAA 121

```

/organism="Silurana tropicalis"

/mol\_type="mRNA"

/db\_xref="taxon:8364"

/clone="TGas123k02"

/dev\_stage="gastrula (stages 10.5-13 mixed)"

/lab\_host="Escherichia coli XL1-blue"

/clone\_lib="XGc-gastrula"

/note="vector: pCS107; Site\_1: EcoRI; Site\_2: NotI; cDNA was oligo dt primed from Sug of poly A+ RNA from stages 10-13 gastrulae. EcoRI-NotI cut cDNA was then ligated into pCS107 with EcoRI at the 5' end and NotI at the 3' end."

BASE COUNT 180 a 166 c 130 g 154 t 1 others

ORIGIN

Query Match 27.4%; Score 130; DB 9; Length 631;

Best Local Similarity 60.8%; Pred. No. 5.6e-19;

Matches 230; Conservative 0; Mismatches 145; Indels 3; Gaps 1;

Qy 1 GCCAGATTATTCAGTGGCAATTTCTTGGTATTATTTGGCCGCAATTTATCATATA 60

Db 80 GCCCGGGATATTCGGTGGACATATTCCTTGGACTGAACCTACTGGCTTTCGACCATC 139

Qy 61 GTTTTTCCTATGGAAGCATGTTTATAGTGTTCATCAAAAGTCCCATACAGCACTGAA 120

Db 140 GTCCTCTCTACAGCAGCATGTTCTACCTCATTAAGACTGGAGCCAAACACAGAGA- 198

Qy 121 ATACGGATCAAGTTTAAAGAGATGATCCCTGCGCAACGTTTTTCTTTATAGTATT 180

Db 199 --ACGACGCGCTGTCAAGAGAGTCCACATGCGAAGAGTCTCTTCTCATAGTTTC 256

Qy 181 ACTGATGATATATGCTGGATACCAATTTTGTAGCGAAAGCTCTTCACTGCTTCAGTA 240

Db 257 ACGATGCTCTCTCTGATGCCATATCTTCTTACTAAGACAGTTCTCTACAGGAGCA 316

Qy 241 GAATACAGGTACCATAACTCTTGGGTAGTATGTTATTTCTGCCATTTAACAGTGCT 300

Db 317 GAGATTCAGGACCATCACTTCATGGATTGTGATCTCATCTCGCTGCTATTAAACAGTCT 376

Qy 301 TTGAACCAATCTCTATCTATCTGACCAAGACCATTTAAAGAAATGATTCATCGTTT 360

Db 377 CTGAACCCCATATTGTACACATCACCACCTCTCTCTTCCAGGAGAGTAACAGTGT 436

Qy 361 TGGCATPACTACAGACAA 378

Db 437 CTACAGAGCAAAAGGCAA 454

RESULT 11

AW436170

LOCUS

DEFINITION 75600 MARC 2PTG Sus scrofa cDNA 5', mRNA linear EST 09-JUL-2000

ACCESSION AW436170

VERSION AW436170.1 GI:6971476

KEYWORDS EST.

SOURCE Sus scrofa (pig)

ORGANISM Sus scrofa

REFERENCE 1 (bases 1 to 235)

AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

TITLE Fahnenkrug, S.C., Smith, T.P.L., Preking, B.A., Cho, J., White, J., J., and Keele, J.W.

Porcine gene discovery by normalized cDNA-library sequencing and EST cluster assembly

JOURNAL Mamm. Genome 13 (8), 475-478 (2002)

MEDLINE 22213789

PUBMED 12226715

COMMENT

Contact: Smith TPL

USDA, ARS, US Meat Animal Research Center

PO Box 166, Clay Center, NE 68933-0166, USA

Tel: 402 762 4366

Fax: 402 762 4390

Email: smithemail.marc.usda.gov

Single pass sequencing. Bases called and trimmed with phred v0.980904.e. Vector identified by cross\_match with the -minscore 20 and -minmatch 12 options.

PCR Primers

FORWARD: AGGAAACAGCTATGACCAT

BACKWARD: GTTTCACAGTCAGCAGC

Plate: 34 row: F column: 22

Seq primer: ATTAGTGACACTATAG.

Location/Qualifiers

1. .235

source

/organism="Sus scrofa"

/mol\_type="mRNA"

/db\_xref="taxon:9823"

/tissue\_type="pooled"

/lab\_host="DH10B"

/clone\_lib="MARC 2PTG"

/note="vector: pCMV SPORT6; Site\_1: NotI; Site\_2: SalI;

Library made from pooled tissue from testis, ovary,

endometrium, hypothalamus, pituitary, and placenta."

BASE COUNT 49 a 64 c 45 g 77 t

ORIGIN

Query Match 20.8%; Score 98.4; DB 9; Length 235;

Best Local Similarity 56.5%; Pred. No. 7.5e-12;

Matches 141; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

Qy 160 CGTTTTCCTTATAGTATTACTGATCATATATGCTGGATACCCATTTTGTAGCGAAA 219

Db 5 CTTTTCCTTATAGGTTCTCTGATGCCATCTCGTGGATTCGCGTATTTGTCATTTAA 64

Qy 220 CCTCTTCACCTGCTCAGGTAGAAATACAGTACCATACCTCTTGGGTAGTATGTTGT 279

Db 65 ATCTCTCCCTCTCTCGGGTGGAAATACAGGCTCCTCCTCTCTGGGGGTCTATTTT 124

Qy 280 TATTCGTGCATTACAGTCTGTTGAACCAATTCCTATATCTGACACAGACCATTTT 339

Db 125 TTCTCTCGGTGACAGTGCCTTGAACCAATCTTTTACACTCTCACAACACAGCTTTTC 184

Qy 340 AAAGAATGATTCATCGTTTGTGGCACTA 371

Db 185 AAGGACAGTTGAACAGTGTGTCACACCA 216

RESULT 12

CA997681

LOCUS

DEFINITION

PIGR8 similar to Porcine LGR8 (INSL3 receptor), mRNA sequence.

CA997681

ACCESSION

CA997681

KEYWORDS

EST.

SOURCE

ORGANISM

Sus scrofa (pig)

Sus scrofa

REFERENCE

1 (bases 1 to 198)

AUTHORS

Bartol, F.F., Wiley, A.A., Bathgate, R.A.D. and Bagnell, C.A.

TITLE

INS3 receptor (LGR8) expression by neonatal porcine uterine

tissues

JOURNAL

Unpublished

COMMENT

Contact: Bartol FF

Department of Animal Sciences

Auburn University, Cellular and Molecular Biosciences Program

240 Upchurch Hall, Auburn, AL 36849-5415, USA

Tel: 334 844 1506

Fax: 334 844 1519

Email: bartolff@auburn.edu

Sequence of cDNA (198 bp) generated by RT-PCR from total RNA

extracted from porcine neonatal uterine tissues obtained on

postnatal day 14.

FEATURES

Location/Qualifiers

1. .198

Db 8 TTTTCCCTATATACATGTCGTCGTCATCAAAAACGCCCTCCAGCACAGAG 67  
 QY 122 TACGGAATCAAGTTAAAGAGAGATGATCCTTGGCAACGTTTCTTTATAGTATTA 181  
 Db 68 TAAGGAATGTTTGGAGAGAGAGTGGCTGTGCAAAATCGTTCTTTTATAGTCTCT 127  
 QY 182 CTGATCATTATGCGGATACCCATTTTGTAGCGAAACCTCTTCTACCTGCTTCAGGTAG 241  
 Db 128 CTGATCCATCTGCGTATCTGATTTGATTTGATTTGATTTGATTTGATTTGATTTG 187  
 QY 242 AATACCAGG 251  
 Db 188 AATACCAG 197

RESULT 15  
 CC244249/c  
 LOCUS CH261-1205.RM1.1 CH261 Gallus gallus genomic clone CH261-1205, 1078 bp DNA linear GSS 12-MAY-2003  
 DEFINITION genomic survey sequence.

ACCESSION CC244249  
 VERSION CC244249.1 GI:30570912  
 KEYWORDS GSS.  
 SOURCE Gallus gallus (chicken)  
 ORGANISM Gallus gallus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
 Phasianinae; Gallus.

REFERENCE 1 (bases 1 to 1078)  
 AUTHORS Kremitzki,C., Higginbotham,J., Wylie,K., Carter,J., McPherson,J., Warren,W., Graves,T., Mardis,E. and Wilson,R.  
 TITLE Gallus gallus BAC End Reads  
 JOURNAL Unpublished  
 COMMENT Contact: Richard K. Wilson  
 Genome Sequencing Center  
 Washington University School of Medicine  
 Email: submissions@watson.wustl.edu  
 Insert Length: 182000 Std Error: 0.00  
 Seq primer: RM1 TACGACTCACTATAGGAGA  
 Class: BAC ends  
 High quality sequence start: 34  
 High quality sequence stop: 532.

FEATURES  
 source  
 1..1078  
 Location/Qualifiers  
 /organism="Gallus gallus"  
 /mol\_type="genomic DNA"  
 /strain="Red Jungle Fowl"  
 /db\_xref="taxon:9031"  
 /clone="CH261-1205"  
 /sex="female"  
 /cell\_line="UCD001, inbred 256"  
 /clone\_lib="CH261"  
 /note="Vector: pTABAC2.1; Site\_1: EcoRI; Site\_2: EcoRI;  
 CH261 Female Chicken library - for library and clone  
 ordering information: http://www.chori.org/bacpac"  
 BASE COUNT 240 a 176 c 141 g 219 t 302 others  
 ORIGIN

Query Match 15.1%; Score 71.8; DB 29; Length 1078;  
 Best Local Similarity 63.6%; Pred. No. 6.8e-06;  
 Matches 143; Conservative 0; Mismatches 77; Indels 5; Gaps 2;  
 QY 31 GGTATTAATTTGGCCGCAATTTATCATCATAGTATTTTCTTATGGAGCATGTTTATAGT 90  
 Db 296 GGTGTGAACCTGATGCTTATCATCATAGTATTTTCTTATGGAGCATGTTTATAGT 237  
 QY 91 GTTCATCAAAATGCCATACAGCACTGAATACGGAATCAAGTTAAAAAGAGATGATC 150  
 Db 236 ATTCAAAAAATCTCTCTCAGGCA--GAAGTGAAGATCAATTCACAGGAGGTACT 180  
 QY 151 TTCTCCAAACGT--TTTTCCTTTATAGTATTTACTGATGATTTGCTGGATACCATTT 208

Db 179 GTTGCCAATCAATATTTTTTTTCTAGTATTCACATGATGCCACCTGCTGGATTCGTCT 120  
 QY 209 TTGTACCGAAACCTCTTTTCTACTGCTTCAGGTAGAAAATACAGGTA 253  
 Db 119 TTGTCAATAAAACCCCTTTCCCTTACTCCAGGAGAAAATGTAGGTA 75

Search completed: September 25, 2003, 14:26:53  
 Job time : 1653 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 25, 2003, 14:59:01 ; Search time 19 seconds  
(without alignments)  
794.657 Million cell updates/sec

File: US-10-049-568-2

Perfect score: 826

Sequence: 1 AQIYSVAIFLGINLAFLII.....HMLHSSGGKCGCHRRCHLS 157

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

PIR76.\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	207.5	25.1	1115	S40241	G protein-coupled
2	143.5	17.4	695	JC1493	folliotropin recept
3	143	17.3	925	JC2033	G protein-coupled
4	141.5	17.1	695	JN0898	folliotropin recept
5	140.5	17.0	696	A41344	lutropin-choriogon
6	139.5	16.9	692	A34548	folliotropin recept
7	139.5	16.9	695	I45896	follicle stimulat
8	136.5	16.5	694	JC4301	folliotropin recept
9	134	16.2	700	I77463	lutinizizing hormon
10	134	16.2	700	A49744	lutropin-choriogon
11	134	16.2	700	A43995	lutropin receptor
12	134	16.2	814	JC7389	thyroid stimulat
13	132	16.0	793	JC7390	thyroid stimulat
14	131.5	15.9	694	JC2337	folliotropin recept
15	131	15.9	699	QRHUT	lutropin-choriogon
16	128	15.5	320	SL1717	probable G protein
17	128	15.5	337	S68678	adenosine receptor
18	127	15.4	764	JC5643	thyroid stimulat
19	125.5	15.2	695	QRHUT	folliotropin recept
20	124	15.0	696	JC7361	folliotropin recept
21	124	15.0	764	A35956	folliotropin recept
22	122	14.8	764	A40077	thyrotropin recept
23	121	14.6	764	I48882	thyrotropin recept
24	119.5	14.5	332	A42171	A2-adenosine recep
25	119.5	14.5	378	JC5245	G protein-coupled
26	119.5	14.5	417	E30341	alpha-1-adrenergic
27	119.5	14.5	517	A45121	alpha-1B adrenergic
28	118	14.3	332	I48933	adenosine receptor
29	117.5	14.2	515	A40491	alpha-1-adrenergic

30	117.5	14.2	764	1	QRHURH	thyrotropin recept
31	117	14.2	326	2	I48096	A1 adenosine recep
32	117	14.2	326	2	A38144	adenosine receptor
33	117	14.2	332	2	JC1229	adenosine receptor
34	115	13.9	326	2	A53005	adenosine receptor
35	114	13.8	466	2	T26458	hypothetical prote
36	113.5	13.7	320	2	A46152	A3 adenosine recep
37	113.5	13.7	392	2	S65693	opioid receptor mu
38	113.5	13.7	400	2	I56553	mu opiate receptor
39	113	13.7	326	2	C30341	G protein-coupled
40	113	13.7	444	2	T27866	hypothetical prote
41	112	13.6	328	2	JN0675	adenosine receptor
42	111	13.4	185	2	I48931	adenosine receptor
43	111	13.4	326	2	A40376	adenosine receptor
44	111	13.4	386	1	S34043	oxytocin receptor
45	110.5	13.4	324	2	S55302	A(1) adenosine rec

ALIGNMENTS

RESULT 1

S40241

G protein-coupled receptor - great pond snail

C:Species: Lymnaea stagnalis (great pond snail)

C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 15-Jun-2001

C:Accession: S40241

R:Tensen, C.P.; Kesteren, E.R.; Planta, R.J.; Cox, K.; Burke, J.F.; Heerikhuisen, H.

submitted to the EMBL Data Library, June 1993

A:Description: A G protein-coupled receptor with LDL-binding motifs suggests a role

A:Reference number: S40241

A:Accession: S40241

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-1115 <TN>

A:Cross-references: EMBL Z23104; NID:G438128; PID:G438129

C:Superfamily: great pond snail LDL receptor-related G protein-coupled receptor; LDL

C:Keywords: G protein-coupled receptor; transmembrane protein

F:38-77/Domain: LDL receptor ligand-binding repeat homology <LDL1>

F:79-113/Domain: LDL receptor ligand-binding repeat homology <LDL2>

F:118-153/Domain: LDL receptor ligand-binding repeat homology <LDL3>

F:158-194/Domain: LDL receptor ligand-binding repeat homology <LDL4>

F:195-230/Domain: LDL receptor ligand-binding repeat homology <LDL5>

F:233-267/Domain: LDL receptor ligand-binding repeat homology <LDL6>

F:274-316/Domain: LDL receptor ligand-binding repeat homology <LDL7>

F:322-361/Domain: LDL receptor ligand-binding repeat homology <LDL8>

F:367-401/Domain: LDL receptor ligand-binding repeat homology <LDL9>

F:406-440/Domain: LDL receptor ligand-binding repeat homology <LDL10>

F:446-483/Domain: LDL receptor ligand-binding repeat homology <LDL11>

F:488-523/Domain: LDL receptor ligand-binding repeat homology <LDL12>

F:584-607/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR1>

F:608-631/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR2>

F:632-655/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR3>

F:656-679/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR4>

F:704-727/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR5>

F:774-797/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR6>

Query Match 25.1%; Score 207.5; DB 2; Length 1115;

Best Local Similarity 33.1%; Pred. No. 6.9e-13;

Matches 51; Conservative 25; Mismatches 45; Indels 33; Gaps 5;

QY 4 YSVAIFLGINLAFLIIIVFSYSGMFYSV---HOSAITATEIRNOVKKEMILAKRFFFTVF 60

||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|

Db 938 YSNGVFILLNLSFVLIIASSYLWMF-SVAKKTRSAVTAESKN---DNAMARRMTLIV 992

QY 61 TDALCWIFPVAKPLSLLQVEIPGTTTSWVIGYSAINSLNPILYTLTRPF----- 113

|| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|

Db 993 TDFCCWPIIVLGVFSLAGARADDQYIAIVFLPNSATNPVYTLSTAPFLGNVKR 1052

QY 114 ----KEMTHRF-----WHNYRQKS 129

|| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|

Db 1053 ANFRKRSFIHSFTGDKTSYVDGTTHTSYCEKKS 1086

## RESULT 2

JC1493  
 N:folllitropin receptor - sheep  
 C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)  
 C:Date: 03-Feb-1994 #sequence\_revision 03-Feb-1994 #text\_change 21-Jan-2000  
 C:Accession: JC1493; I47080  
 R:Khan, H.; Yarney, T.A.; Sairam, M.R.  
 Biochem. Biophys. Res. Commun. 190, 888-894, 1993  
 A:Title: Cloning of alternately spliced mRNA transcripts coding for variants of ovine testis  
 A:Reference number: JC1493; MUID:93176195; PMID:8439338  
 A:Accession: JC1493  
 A:Molecule type: mRNA  
 A:Residues: 1-695 <XHA>

A:Experimental source: testis  
 R:Yarney, T.A.; Sairam, M.R.; Khan, H.; Ravindranath, N.; Payne, S.; Seidah, N.G.  
 Mol. Cell. Endocrinol. 93, 219-226, 1993  
 A:Title: Molecular cloning and expression of the ovine testicular follicle stimulating hormone  
 A:Reference number: I47080; MUID:93351750; PMID:8394255  
 A:Accession: I47080  
 A>Status: preliminary; translated from GE/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-695 <VAR>  
 A:CROSS-references: GB:I470302; NID:g165884; PIDN:AAA31525.1; PID:g165885

C:Genetics:  
 A:Gene: FSH-R  
 C:Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repeat  
 C:Keywords: G protein-coupled receptor; glycoprotein; transmembrane protein  
 F:1-95/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR>  
 F:191-199/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 17.4%; Score 143.5; DB 2; Length 695;  
 Best Local Similarity 27.2%; Pred. No. 1.4e-06;  
 Matches 44; Conservative 32; Mismatches 59; Indels 27; Gaps 6;

QY 1 AQIYSVAIFGLINLAFAFIIVFSYGMFYSVHQSATATATIRNOVKEMILAKRFFIYFV 60

Db 527 SOLYVMSL-LVLNLVAFWICGCTHYLTVRNPVSS-----SSDTAKRAMLIF 579

QY 61 TDALCWPII-FVAKPLSLQVEIGTITSVWVIGYSAINSLNPLILYTLTRPFKE- 115

Db 580 TDFLCMAPISFFAISAKLVPLITVSKSKILLVLFVFPINSCANFLYAITRNRDFFI 639

QY 116 MIHFW-----HNTRQKSKMDSKGRKMKLHHSSGKGHC 151

Db 640 LLSKFGCYEQVQATYRSTFTA-----HNEHPRNGHC 672

## RESULT 3

JC2033  
 G protein-coupled receptor - sea anemone (Anthopleura elegantissima)  
 C:Species: Anthopleura elegantissima  
 C:Date: 30-Sep-1993 #sequence\_revision 20-Aug-1994 #text\_change 21-Jul-2000  
 C:Accession: JC2033; S41908

R:Nothacker, H.P.; Grimmelikhuijzen, C.J.P.  
 Biochem. Biophys. Res. Commun. 197, 1062-1069, 1993  
 A:Title: Molecular cloning of a novel, putative G protein-coupled receptor from sea anemone  
 A:Reference number: JC2033; MUID:94107299; PMID:8280121  
 A:Accession: JC2033

A:Molecule type: mRNA  
 A:Residues: 1-925 <NOT>

A:CROSS-references: EMBL:228332; NID:g452810; PIDN:CAA82186.1; PID:g452811

C:Superfamily: leucine-rich alpha-2-glycoprotein repeat homology

C:Keywords: G protein-coupled receptor; glycoprotein; neuropeptide; phosphoprotein; transmembrane protein

F:61-84/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>  
 F:85-109/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>  
 F:110-133/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>  
 F:134-157/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>  
 F:333-349,350-384,385-419,420-453,454-462/Region: glycine-rich

F:330-551/Domain: transmembrane #status predicted <TM1>  
 F:562-584/Domain: transmembrane #status predicted <TM2>  
 F:607-628/Domain: transmembrane #status predicted <TM3>

F:652-673/Domain: transmembrane #status predicted <TM4>  
 F:692-712/Domain: transmembrane #status predicted <TM5>  
 F:740-763/Domain: transmembrane #status predicted <TM6>  
 F:775-795/Domain: transmembrane #status predicted <TM7>  
 F:61-152,212,435,491,495/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 F:553,645,855,859,868,894,900/Binding site: phosphate (Ser) (covalent) #status predicted  
 F:649,853/Binding site: phosphate (Thr) (covalent) #status predicted

Query Match 17.3%; Score 143; DB 2; Length 925;

Best Local Similarity 33.3%; Pred. No. 2.1e-06;  
 Matches 37; Conservative 21; Mismatches 47; Indels 6; Gaps 3;

QY 6 VAIFGLINLAFAFIIVFSYGMFYSVHQSATATATIRNOVK-EMILAKRFFIYFVTDAL 64

Db 694 VAPLLLVNGASPIVMYLSRLVYV---VSGDMGAPKRNDSKVAKRMALIVFTDML 749

QY 65 CWIPI-FVAKPLSLQVEIGTITSVWVIGYSAINSLNPLILYTLTRPFK 114

Db 750 CWAPTAFFGLLAAGQTLVNTQSKILLVFPFINSICNPFYAFFTKAFK 800

## RESULT 4

JN0898

folllitropin receptor precursor - crab-eating macaque

N:Alternate names: follicle-stimulating hormone receptor (FSHR)

C:Species: Macaca fascicularis (crab-eating macaque)

C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999

C:Accession: JN0898; S36452

R:Gromoll, J.; Dankbar, B.; Sharma, R.S.; Nieschlag, E.

Biochem. Biophys. Res. Commun. 196, 1066-1072, 1993

A:Title: Molecular cloning of the testicular follicle stimulating hormone receptor of

A:Reference number: JN0898; MUID:94071854; PMID:7504463

A:Accession: JN0898

A:Molecule type: mRNA

A:Residues: 1-695 <GRO>

A:CROSS-references: EMBL:X74454; NID:g396801; PIDN:CA852463.1; PID:g396802

A:Note: the authors translated the codon AGT for residue 488 as Arg

C:Function:

A:Description: receptor that mediates the biochemical effects of follitropin

C:Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repeat

C:Keywords: G protein-coupled receptor; glycoprotein; hormone receptor; phosphoprotein

F:1-17/Domain: signal sequence #status predicted <SIG>

F:18-695/Product: follitropin receptor #status predicted <PR>

F:71-95/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>

F:96-120/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>

F:121-145/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>

F:146-169/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>

F:172-193/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>

F:194-218/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>

F:367-387/Domain: transmembrane #status predicted <TM1>

F:399-421/Domain: transmembrane #status predicted <TM2>

F:444-465/Domain: transmembrane #status predicted <TM3>

F:486-508/Domain: transmembrane #status predicted <TM4>

F:529-550/Domain: transmembrane #status predicted <TM5>

F:574-597/Domain: transmembrane #status predicted <TM6>

F:609-630/Domain: transmembrane #status predicted <TM7>

F:191,199,293,318/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:555/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicted

F:596/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status predicted

Query Match 17.1%; Score 141.5; DB 1; Length 695;

Best Local Similarity 26.5%; Pred. No. 2.2e-06;

Matches 44; Conservative 32; Mismatches 63; Indels 27; Gaps 6;

QY 1 AQIYSVAIFGLINLAFAFIIVFSYGMFYSVHQSATATATIRNOVKEMILAKRFFIYFV 60

Db 527 SOLYVMSL-LVLNLVAFWICGCTHYLTVRNPVSS-----SSDTAKRAMLIF 579

QY 61 TDALCWPII-FVAKPLSLQVEIGTITSVWVIGYSAINSLNPLILYTLTRPFKE- 115

Db 580 TDFLCMAPISFFAISAKLVPLITVSKAKILGLVLEFPINSCANFLYAITRNRDFFI 639

QY 116 MIHFW-----HNTRQKSKMDSKGRKMKLHHSSGKGHC 155



## TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-628-5197

TELEFAX: 202-737-3528

TELEX: 248633

INFORMATION FOR SEQ ID NO: 18:

SEQUENCE CHARACTERISTICS:

LENGTH: 370 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

PCT-US93-08528-18

## Query Match

Best Local Similarity 16.5%; Score 136; DB 5; Length 370;

Matches 39; Conservative 24; Mismatches 50; Indels 24; Gaps 5;

QY 2 QIYSVAIFLGINLAFFIIIVFSYSGMSFYVS-----HQAATATATIRNOVKKEMILAK 53  
Db 178 RVIIVAKRTTKNLEAGVVKMS-NSKELTLRIHWSKNFHEDTLSSTRAKGNPRSSIAVK 236  
QY 54 RFFP-----IVFTDALCWIPFIIVAKPLSLI--QVEIPGTTTSWVWIGYSAINSA 100  
Db 237 LKFSREKKAATLGIIVGVWLQWLPFIALPLGSLFSLKPPDAVKKFWLGY--FNSC 294  
QY 101 LNPILYTLTRPFKMI 117  
Db 295 LNPFIYPCSKFEKRAL 311

## RESULT 7

US-07-757-342D-3

Sequence 3, Application US/07757342D

Patent No. 6218509

GENERAL INFORMATION:

APPLICANT: IGARASHI, Masao

MINEGISHI, Takashi

NAKAMURA, Kazuo

TITLE OF INVENTION: PROTEIN, DNA AND USE THEREOF

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:

ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &amp;

CUSHMAN

STREET: 130 Water Street

CITY: Boston

STATE: Massachusetts

COUNTRY: US

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/757,342D

FILING DATE: 10-Sep-1991

CLASSIFICATION: &lt;Unknown&gt;

ATTORNEY/AGENT INFORMATION:

NAME: BUCKLEY, Linda M.

REGISTRATION NUMBER: 31003

REFERENCE/DOCKET NUMBER: 41226

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617)523-3400

TELEFAX: (617)523-6440

TELEX: 200291 STRE UR

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 700 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 3:

US-07-757-342D-3

Query Match 16.2%; Score 134; DB 3; Length 700;  
Best Local Similarity 27.2%; Pred. No. 2.9e-06;  
Matches 34; Conservative 31; Mismatches 48; Indels 12; Gaps 4;  
QY 1 AQIYSVAIFLGINLAFFIIIVFSYSGMSFYVSHQSATATATIRNOVKKEMILAKRFFPIVF 60  
Db 528 SQVYILSILI-LNVVAFVVICACIRIYFAVONPELTAP-----NKDTKIAKKMAILIF 580  
QY 61 TDALCWIPFI-FVAKPLSLQLQVEIPGTTTSWVWIGYSAINSAALNPILYTLTRPKE--- 115  
Db 581 TDFTCMAPISFFAISAAFKVPLIIVTNSKILLVLPVNSCANPFLYAIPTKAFQDFLL 640  
QY 116 MIHRF 120  
Db 641 LLSRF 645

## RESULT 8

US-07-757-342D-8

Sequence 8, Application US/07757342D

Patent No. 6218509

GENERAL INFORMATION:

APPLICANT: IGARASHI, Masao

MINEGISHI, Takashi

NAKAMURA, Kazuo

TITLE OF INVENTION: PROTEIN, DNA AND USE THEREOF

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:

ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &amp;

CUSHMAN

STREET: 130 Water Street

CITY: Boston

STATE: Massachusetts

COUNTRY: US

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/757,342D

FILING DATE: 10-Sep-1991

CLASSIFICATION: &lt;Unknown&gt;

ATTORNEY/AGENT INFORMATION:

NAME: BUCKLEY, Linda M.

REGISTRATION NUMBER: 31003

REFERENCE/DOCKET NUMBER: 41226

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617)523-3400

TELEFAX: (617)523-6440

TELEX: 200291 STRE UR

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 611 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 8:

US-07-757-342D-8

## Query Match

Best Local Similarity 15.9%; Score 131; DB 3; Length 611;

Matches 34; Conservative 25; Mismatches 48; Indels 8; Gaps 3;

QY 1 AQIYSVAIFLGINLAFFIIIVFSYSGMSFYVSHQSATATATIRNOVKKEMILAKRFFPIVF 60  
Db 436 SQVYILSILI-LNVVAFVVICACIRIYFAVONPELTAP-----NKDTKIAKKMAILIF 488  
QY 61 TDALCWIPFI-FVAKPLSLQLQVEIPGTTTSWVWIGYSAINSAALNPILYTLTRPKE 114  
Db 489 TDFTCMAPISFFAISAAFKVPLIIVTNSKILLVLPVNSCANPFLYAIPTKTFQ 543

```

QY      116 MIHRF 120
      ::||
Db      641 LLSRF 645

RESULT 12
JC7389
Thyroid stimulating hormone receptor a - salmon
N:Alternate names: thyrotropin receptor a
C:Species: Oncorhynchus sp. (salmon)
C:Date: 17-Nov-2000 #sequence_revision 17-Nov-2000 #text_change 07-Mar-2003
C:Accession: JC7389
R:Oba Y.; Hirai, T.; Yoshiura, Y.; Kobayashi, T.; Nagahama, Y.
Biochem. Biophys. Res. Commun. 276, 289-263, 2000
A:Title: Cloning, functional characterization, and expression of thyrotropin receptor
A:Reference number: JC7389
A:Contents: Thyroid
A:Accession: JC7389
A:Molecule type: mRNA
A:Residues: 1-814 <OBA>
A:Cross-references: DDBJ:AB030954
C:Comment: This receptor, a transmembrane protein, which mediates the actions of thy
C:Genetics:
A:Gene: tsh-ra
A:Introns: 610/3
C:Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repe
C:Keywords: glycoprotein; hormone receptor; thyroid gland; transmembrane protein

Query Match      16.2%; Score 134; DB 2; Length 814;
Best Local Similarity 26.8%; Pred. No. 1.5e-05;
Matches 44; Conservative 32; Mismatches 60; Indels 28; Gaps 6;

QY      1 AQIYSAIFLGINLAAFIIVFSYGSMTFVSQSAITATEIRNOYKKEMILAKREFFIVF 60
      |||:|::: |:|:|::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
Db      624 AQVIIVSLI-LNLIAFLVLCACIKYICVHPHYRSG-----SKDTNIAKMAVLIF 676

QY      61 TDALCWIP I-FVAKPLSLQVPEITPTTSWVIGYSAINSALNPILYTLTTRPKF----- 114
      ||| ||| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
Db      677 TDFLCMAPIFYAMSVAIDLRLITVSNKLLVFLYPLNSCANPELYAIFTKAFRGDVF I 736

QY      115 -----EMIHFWENYEQKSDSKGIRKXHLHSSGKGKCGHCR 153
      ||| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
Db      737 LLSKVLGICQRRALFR-GQTVSSKG-----SSGVC HQGR 770

RESULT 13
JC7390
Thyroid stimulating hormone receptor b - salmon
N:Alternate names: thyrotropin receptor b
C:Species: Oncorhynchus sp. (salmon)
C:Date: 17-Nov-2000 #sequence_revision 17-Nov-2000 #text_change 07-Mar-2003
C:Accession: JC7390
R:Oba, Y.; Hirai, T.; Yoshiura, Y.; Kobayashi, T.; Nagahama, Y.
Biochem. Biophys. Res. Commun. 276, 258-263, 2000
A:Title: Cloning, functional characterization, and expression of thyrotropin recepto
A:Reference number: JC7389
A:Contents: Thyroid
A:Accession: JC7390
A:Molecule type: mRNA
A:Residues: 1-793 <OBA>
A:Cross-references: DDBJ:AB030955
C:Comment: This receptor, a transmembrane protein, which mediates the actions of thy
C:Genetics:
A:Gene: tsh-rb
A:Introns: 584/3
C:Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repe
C:Keywords: glycoprotein; hormone receptor; thyroid gland; transmembrane protein

Query Match      16.0%; Score 132; DB 2; Length 793;
Best Local Similarity 27.3%; Pred. No. 2.4e-05;
Matches 33; Conservative 28; Mismatches 40; Indels 20; Gaps 4;

```





## RESULT 14

US-08-795-876-2  
 ; Sequence 2, Application US/08795876  
 ; Patent No. 6403305  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gershengorn, Marvin C.  
 ; APPLICANT: Geras-Raaka, Elizabeth  
 ; APPLICANT: Nussenzeveig, Daniel R.  
 ; TITLE OF INVENTION: STRATEGY TO CLONE DRUGS FOR G PROTEIN  
 ; TITLE OF INVENTION: COUPLED RECEPTORS  
 ; NUMBER OF SEQUENCES: 44  
 ; CORRESPONDENCE ADDRESSES:  
 ; ADDRESSEE: NIXON, HARGRAVE, DEVANS & DOYLE LLP  
 ; STREET: Clinton Square, P.O. Box 1051  
 ; CITY: Rochester  
 ; STATE: New York  
 ; COUNTRY: USA  
 ; ZIP: 14603  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/795,876  
 ; FILING DATE:  
 ; CLASSIFICATION: 530  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: BRAMAN, SUSAN J.  
 ; REGISTRATION NUMBER: 34,103  
 ; REFERENCE/DOCKET NUMBER: 19603/1280  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 716-263-1636  
 ; TELEFAX: 716-263-1600  
 ; INFORMATION FOR SEQ ID NO: 2:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 436 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: not relevant  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: peptide  
 ; US-08-795-876-2

Query Match 15.2%; Score 125.5; DB 4; Length 436;  
 Best Local Similarity 26.1%; Pred. No. 1.4e-05;  
 Matches 43; Conservative 30; Mismatches 59; Indels 33; Gaps 6;  
 QY 1 AQIYSVAIFGLINLAIFIIVESYGSFYSVHQSATATEIRNQVKEMILAKREFFIYF 60  
 Db 268 SOLYVMSL-LVLNLAFFVVICGYIHLYLTVRNPIVSS-----SSDTIAKRMAMLI 320  
 QY 61 TDALCWPI-FVAKPISLIQVLPCTITTSWVIGYSAINSLNPILYILTRPK----- 114  
 Db 321 TDFLCMAPIFFFAISASLVPLITVSKAKILLVLFHPINSCANPELYAIFTKNERDFFI 380  
 QY 115 -----EMHFRFWNYRQRKSDSKGIRKMHLMHSSGGKGC 151  
 Db 381 LLSKCGCYEMQAOI---YTEITSSTV-----HNTHPRNGHC 413

## RESULT 15

US-08-487-886-2  
 ; Sequence 2, Application US/08487886  
 ; Patent No. 5744448  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Kelton, Christie Ann  
 ; APPLICANT: Schweickhardt, Rene Lynn  
 ; APPLICANT: Cheng, Shirley Vui Yen  
 ; APPLICANT: Nugent, No. 3744448een Patrice  
 ; TITLE OF INVENTION: Human Follicle Stimulating  
 ; TITLE OF INVENTION: Hormone Receptor  
 ; NUMBER OF SEQUENCES: 2

; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Stephan P. Williams,  
 ; ADDRESSEE: Ares-Serono, Inc.  
 ; STREET: Exchange Place, 37th floor  
 ; CITY: Boston  
 ; STATE: MA  
 ; COUNTRY: USA  
 ; ZIP: 02109  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: 3.5" diskette, 1.44 MB, high density  
 ; COMPUTER: IBM PS/2, model 55 SX  
 ; OPERATING SYSTEM: MS-DOS version 4.0  
 ; SOFTWARE: VAX/VMS Massll via Kermit to IBM MS-DOS  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/487,886  
 ; FILING DATE: 07-JUN-1995  
 ; CLASSIFICATION: 514  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 07/670,085  
 ; FILING DATE: 15-MAR-1991  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Williams, Stephan P.  
 ; REGISTRATION NUMBER: 28546  
 ; REFERENCE/DOCKET NUMBER: US/252  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (617) 723-1300  
 ; TELEFAX: (617) 723-8923  
 ; INFORMATION FOR SEQ ID NO: 2:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 695  
 ; TYPE: Amino acid  
 ; TOPOLOGY: Linear  
 ; MOLECULE TYPE: protein  
 ; FEATURE:  
 ; NAME/KEY: signal sequence  
 ; LOCATION: -17 to -1  
 ; IDENTIFICATION METHOD: hydrophobic  
 ; FEATURE:  
 ; NAME/KEY: putative amino-terminal extracellular domain  
 ; LOCATION: 1 to 349  
 ; IDENTIFICATION METHOD: similarity with other  
 ; IDENTIFICATION METHOD: dimeric glycoprotein receptor extracellular  
 ; IDENTIFICATION METHOD: domains, hydrophilic  
 ; FEATURE:  
 ; NAME/KEY: transmembrane domain  
 ; LOCATION: 350 to 613  
 ; IDENTIFICATION METHOD: similarity to other G  
 ; IDENTIFICATION METHOD: protein-coupled receptor transmembrane domains  
 ; FEATURE:  
 ; NAME/KEY: putative transmembrane region I  
 ; LOCATION: 350 to 370  
 ; IDENTIFICATION METHOD: similarity to other G  
 ; IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,  
 ; IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length  
 ; FEATURE:  
 ; NAME/KEY: putative transmembrane region II  
 ; LOCATION: 382 to 404  
 ; IDENTIFICATION METHOD: similarity to other G  
 ; IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,  
 ; IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length  
 ; FEATURE:  
 ; NAME/KEY: putative transmembrane region III  
 ; LOCATION: 427 to 448  
 ; IDENTIFICATION METHOD: similarity to other G  
 ; IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,  
 ; IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length  
 ; FEATURE:  
 ; NAME/KEY: putative transmembrane region IV  
 ; LOCATION: 469 to 491  
 ; IDENTIFICATION METHOD: similarity to other G  
 ; IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,  
 ; IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length  
 ; FEATURE:



GenCore version 5.1.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 25, 2003, 15:00:41 : Search time 64 Seconds  
(without alignments)  
371.175 Million cell updates/sec

Title: US-10-049-568-2  
Perfect score: 326  
Sequence: 1 AQIYSVAIFLGINLAARFIIT.....HMLHSSGGKGGHRCRCHLS 157

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5  
Searched: 566894 seqs, 151307093 residues  
Total number of hits satisfying chosen parameters: 566894

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published\_Applications\_AA: \*  
1: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep: \*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep: \*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep: \*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep: \*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep: \*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep: \*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep: \*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep: \*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep: \*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep: \*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep: \*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep: \*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep: \*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep: \*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep: \*  
16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep: \*  
17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep: \*  
18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	638	77.2	188	15	US-10-073-885-69
2	638	77.2	396	10	US-09-895-686-4
3	638	77.2	757	10	US-09-928-175-24
4	638	77.2	757	15	US-10-225-567A-623
5	632	76.5	176	9	US-09-764-853-557
6	632	76.5	176	10	US-09-764-877-2009
7	632	76.5	176	11	US-09-989-442-105
8	632	76.5	176	12	US-09-764-886-52
9	632	76.5	176	15	US-10-073-865-84
10	632	76.5	176	15	US-10-103-313-373
11	632	76.5	176	15	US-10-073-885-96
12	409	49.5	718	10	US-09-928-175-21
13	409	49.5	737	10	US-09-928-175-20
14	409	49.5	737	15	US-10-229-735-3
15	392.5	47.5	355	12	US-10-321-807-18

16	392.5	47.5	646	10	US-09-928-175-13	Sequence 13, Appl
17	392.5	47.5	682	10	US-09-928-175-12	Sequence 12, Appl
18	392.5	47.5	694	10	US-09-928-175-8	Sequence 8, Appl
19	392.5	47.5	713	11	US-09-965-538A-6	Sequence 6, Appl
20	392.5	47.5	718	10	US-09-928-175-3	Sequence 3, Appl
21	392.5	47.5	730	10	US-09-928-175-7	Sequence 7, Appl
22	392.5	47.5	737	11	US-09-965-536A-2	Sequence 2, Appl
23	392.5	47.5	754	10	US-09-928-175-2	Sequence 2, Appl
24	392.5	47.5	754	15	US-10-229-735-1	Sequence 1, Appl
25	392.5	47.5	754	15	US-10-222-668-2	Sequence 2, Appl
26	349	42.3	107	11	US-09-930-312-2	Sequence 2, Appl
27	244	29.5	327	11	US-09-782-974C-66	Sequence 66, Appl
28	244	27.2	334	15	US-10-270-333-21	Sequence 21, Appl
29	244	27.2	722	15	US-10-222-668-4	Sequence 4, Appl
30	212.5	25.7	359	15	US-10-270-333-27	Sequence 27, Appl
31	207.5	25.1	1115	11	US-09-965-538A-10	Sequence 10, Appl
32	207.5	25.1	1115	15	US-10-222-668-3	Sequence 3, Appl
33	139.5	16.9	675	9	US-09-877-804-7	Sequence 7, Appl
34	139.5	16.9	688	11	US-09-965-536A-12	Sequence 12, Appl
35	139.5	16.9	692	9	US-09-877-804-6	Sequence 6, Appl
36	139.5	16.9	692	11	US-09-965-536A-11	Sequence 11, Appl
37	134	16.2	674	9	US-09-877-804-3	Sequence 3, Appl
38	134	16.2	700	9	US-09-877-804-2	Sequence 2, Appl
39	131	15.9	458	9	US-09-862-767A-5	Sequence 5, Appl
40	131	15.9	693	11	US-09-965-536A-14	Sequence 14, Appl
41	131	15.9	699	10	US-09-804-626-6	Sequence 6, Appl
42	131	15.9	699	15	US-10-225-567A-150	Sequence 150, App
43	130	15.7	861	9	US-09-804-551B-20	Sequence 20, Appl
44	129.5	15.7	695	12	US-10-349-838A-30	Sequence 30, Appl
45	129	15.6	687	11	US-09-965-536A-13	Sequence 13, Appl

ALIGNMENTS

RESULT 1  
US-10-073-885-69  
; Sequence 69, Application US/10073885  
; Publication No. US20030096346A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: P0303C1  
; CURRENT APPLICATION NUMBER: US/10/073.885  
; CURRENT FILING DATE: 2002-02-14  
; Prior Application removed - See file Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 116  
; SOFTWARE: PatentIn ver. 2.0  
; SEQ ID NO 69  
; TYPE: PRT  
; LENGTH: 188  
; ORGANISM: Homo sapiens  
US-10-073-885-69

Query Match 77.2%; Score 638; DB 15; Length 188;  
Best Local Similarity 94.8%; Pred. No. 2.le-c2;  
Matches 127; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY	1	AQIYSVAIFLGINLAARFIITVSQSMFYSHQSATATATEIRNOVKEMILAKREFPIVF	60
DB	7	AQIYSVAIFLGINLAARFIITVSQSMFYSHQSATATATEIRNOVKEMILAKREFPIVF	66
QY	61	TDALCWIPFVAKPLSLQVEIPGTITSWVIGVYSAINALNPILYTLTTPPFKEMIHRT	120
DB	67	TDALCWIPFVAKPLSLQVEIPGTITSWVIFILPINSALNPILYTLTTPPFKEMIHRT	126
QY	121	WHNYRQKSMDSKG 134	
DB	127	WHNYRQKSMDSKG 140	

RESULT 2  
US-09-895-686-4

```

; Sequence 4, Application US/09895686
; Patent No. US20020106655A1
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Lal, Freeti
; APPLICANT: Tang, Y. Tom
; APPLICANT: Baughn, Mariah R.
; TITLE OF INVENTION: HUMAN GPCR PROTEINS
; FILE REFERENCE: PG-0044 CIP
; CURRENT APPLICATION NUMBER: US/09/895,686
; CURRENT FILING DATE: 2001-06-28
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PERL Program
; SEQ ID NO 4
; LENGTH: 396
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20020106655A1 2488822CD1
US-09-895-686-4

Query Match      77.2%; Score 638; DB 10; Length 396;
Best Local Similarity 94.8%; Pred. No. 5.4e-62;
Matches 127; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 1 AQIYSVAIFLGINLAARFIIIVSYGSMFYSHQSAITATEIRNOVKEMILAKRFFFIYF 60
Db 215 AQIYSVAIFLGINLAARFIIIVSYGSMFYSHQSAITATEIRNOVKEMILAKRFFFIYF 274

Qy 61 TDALCWIPFVAKPLSLQLOVEIPGTTISWVYVIFILPINSALNPILYTLTTRPFKEMIHFR 120
Db 275 TDALCWIPFVAKPLSLQLOVEIPGTTISWVYVIFILPINSALNPILYTLTTRPFKEMIHFR 334

Qy 121 WHNYRQKSMDSKG 134
Db 335 WHNYRQKSMDSKG 348

RESULT 3
US-09-928-175-24
; Sequence 24, Application US/09928175
; Patent No. US20020123618A1
; GENERAL INFORMATION:
; APPLICANT: Paszty, Christopher J.
; APPLICANT: Gong, Jianhua
; APPLICANT: Daugherty, Betsy
; APPLICANT: Rogers, No. US20020123618Alma
; TITLE OF INVENTION: Leucine-Rich G Protein Coupled Receptor-8 Molecules and
; FILE REFERENCE: 00-1229
; CURRENT APPLICATION NUMBER: US/09/928,175
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/224,455
; PRIOR FILING DATE: 2000-08-10
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 24
; LENGTH: 757
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-928-175-24

Query Match      77.2%; Score 638; DB 10; Length 757;
Best Local Similarity 94.8%; Pred. No. 1.2e-61;
Matches 127; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 1 AQIYSVAIFLGINLAARFIIIVSYGSMFYSHQSAITATEIRNOVKEMILAKRFFFIYF 60
Db 576 AQIYSVAIFLGINLAARFIIIVSYGSMFYSHQSAITATEIRNOVKEMILAKRFFFIYF 635

Qy 61 TDALCWIPFVAKPLSLQLOVEIPGTTISWVYVIFILPINSALNPILYTLTTRPFKEMIHFR 120
Db 61 TDALCWIPFVAKPLSLQLOVEIPGTTISWVYVIFILPINSALNPILYTLTTRPFKEMIHFR 120

Qy 121 WHNYRQKSMDSKG 134
Db 696 WHNYRQKSMDSKG 709

RESULT 5
US-09-764-853-557
; Sequence 557, Application US/09764853
; Patent No. US20020090672A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PJ206
; CURRENT APPLICATION NUMBER: US/09/764,853
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 939
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 557
; LENGTH: 176
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (133)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-853-557

Query Match      76.5%; Score 632; DB 9; Length 176;
Best Local Similarity 94.0%; Pred. No. 9e-62;
Matches 136; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

```

```

Db 636 TDALCWIPFVAKPLSLQLOVEIPGTTISWVYVIFILPINSALNPILYTLTTRPFKEMIHFR 695
Qy 121 WHNYRQKSMDSKG 134
Db 696 WHNYRQKSMDSKG 709

RESULT 4
US-10-225-567A-623
; Sequence 623, Application US/10225567A
; Publication No. US20030113798A1
; GENERAL INFORMATION:
; APPLICANT: Lifespan Biosciences
; APPLICANT: Brown, Joseph P.
; APPLICANT: Burner, Glenna C.
; APPLICANT: Roush, Christine L.
; TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTOR
; FILE REFERENCE: 1920-4-4
; CURRENT APPLICATION NUMBER: US/10/225,567A
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/257,144
; PRIOR FILING DATE: 2000-12-19
; NUMBER OF SEQ ID NOS: 2292
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 623
; LENGTH: 757
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-225-567A-623

Query Match      77.2%; Score 638; DB 15; Length 757;
Best Local Similarity 94.8%; Pred. No. 1.2e-61;
Matches 127; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 1 AQIYSVAIFLGINLAARFIIIVSYGSMFYSHQSAITATEIRNOVKEMILAKRFFFIYF 60
Db 576 AQIYSVAIFLGINLAARFIIIVSYGSMFYSHQSAITATEIRNOVKEMILAKRFFFIYF 635

Qy 61 TDALCWIPFVAKPLSLQLOVEIPGTTISWVYVIFILPINSALNPILYTLTTRPFKEMIHFR 120
Db 636 TDALCWIPFVAKPLSLQLOVEIPGTTISWVYVIFILPINSALNPILYTLTTRPFKEMIHFR 695

Qy 121 WHNYRQKSMDSKG 134
Db 696 WHNYRQKSMDSKG 709

```

A;Description: receptor that mediates the biochemical effects of follitropin  
C;Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repeat  
C;Keywords: alternative splicing; G protein-coupled receptor; glycoprotein; hormone  
F;1-15/Domain: signal sequence #status predicted<SIG>  
F;16-692/Product: follitropin receptor #status predicted<MAT>  
F;16-366/Domain: extracellular hormone binding #status predicted<BHB>  
F;56-70/Domain: leucine-rich alpha-2-glycoprotein repeat homology<LRR1>  
F;71-95/Domain: leucine-rich alpha-2-glycoprotein repeat homology<LRR2>  
F;96-120/Domain: leucine-rich alpha-2-glycoprotein repeat homology<LRR3>  
F;121-145/Domain: leucine-rich alpha-2-glycoprotein repeat homology<LRR4>  
F;146-169/Domain: leucine-rich alpha-2-glycoprotein repeat homology<LRR5>  
F;172-193/Domain: leucine-rich alpha-2-glycoprotein repeat homology<LRR6>  
F;194-218/Domain: leucine-rich alpha-2-glycoprotein repeat homology<LRR7>  
F;367-387/Domain: transmembrane #status predicted<TM1>  
F;398-421/Domain: transmembrane #status predicted<TM2>  
F;443-465/Domain: transmembrane #status predicted<TM3>  
F;486-508/Domain: transmembrane #status predicted<TM4>  
F;529-550/Domain: transmembrane #status predicted<TM5>  
F;574-597/Domain: transmembrane #status predicted<TM6>  
F;609-630/Domain: transmembrane #status predicted<TM7>  
F;191\_199\_293/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F;554/Binding site: phosphate (Phe) (covalent) (by protein kinase C) #status predicted  
F;595/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status predicted

Query Match 16.9%; Score 139.5; DB 2; Length 692;  
Best Local Similarity 27.3%; Pred. No. 3.6e-06;  
Matches 47; Conservative 30; Mismatches 60; Indels 35; Gaps 6;

QY 1 AQIYSVAIFLGINLAFTIIIVFSVGSFMFSVHQSATATEIRNOVKKEMILAKRFFFIIVF 60  
:::|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:  
Db 526 SQLYVMAL-LVLNVLAFVVICGCTHYLTVRNPITVSS-----SSDTKIARKMAMLIIF 578

QY 61 TDALCWPII-FYAKPLSLLOVEIPGTITSWYIGVSAINSLPNILYTLTPPFE---- 115  
:::|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:  
Db 579 TDFLCMAPISFAISASLKVPILTIVSRKLLIVLPINSNCANPFLYAIFTNFRDRFFI 638

QY 116 MIHRF-----WHNYQRKSDSKGRKH-----MLRHSS 144  
:::|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:  
Db 639 LLSKFGCYENQAQIYRTTSATHNFARKSHCSSAPRVTVSYVLPLNHSS 690

RESULT 7  
I45896  
follicle stimulating hormone receptor - bovine  
C:Species: Bos primigenius taurus (cattle)  
C>Date: 15-Oct-1996 #sequence\_revision 15-Oct-1996 #text\_change 21-Jan-2000  
C:Accession: I45896  
R;Route, A.; Lambert, A.; Saumande, J.; Silversides, D.W.; Lussier, J.G.  
Mol. Reprod. Dev. 39, 127-135, 1994  
A;Title: Structure of the bovine follicle-stimulating hormone receptor complementary  
A;Reference number: I45896; WUID:95127199; PMID:7826612  
A;Accession: I45896  
A>Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: mRNA  
A;Residues: 1-695 <HOU>  
A;Cross-references: GB:I22319; NID:g404671; PIDN:AAC37324.1; PID:g404672  
C:Genetics:  
A;Gene: FSHR

C;Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repeat  
F;71-95/Domain: leucine-rich alpha-2-glycoprotein repeat homology<LRR>

Query Match 16.9%; Score 139.5; DB 2; Length 695;  
Best Local Similarity 26.5%; Pred. No. 3.6e-06;  
Matches 43; Conservative 33; Mismatches 59; Indels 27; Gaps 6;

QY 1 AQIYSVAIFLGINLAFTIIIVFSVGSFMFSVHQSATATEIRNOVKKEMILAKRFFFIIVF 60  
:::|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:~::~:|:|:|:|:|:  
Db 527 SQLYVMSL-LVLNVLAFVVICGCTHYLTVRNPITSS-----SSDTKIARKMAMLIIF 579

QY 61 TDALCWPII-FYAKPLSLLOVEIPGTITSWYIGVSAINSLPNILYTLTPPFE---- 115  
:::|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:~::~:|:|:|:|:|:  
Db 580 TDFLCMAPISFAISASLKVPILTIVSRKLLIVLPINSNCANPFLYAIFTNFRDRFFI 639

F;153-177/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRP>

Query Match 16.2%; Score 134; DB 2; Length 700;  
Best Local Similarity 27.2%; Pred. No. 1.3e-05;  
Matches 34; Conservative 31; Mismatches 48; Indels 12; Gaps 4;

QY 1 AQIYSVAIPGLINLAAPFIIVFSYGSMFYSVHQSAITATETNRNOVKKEMILAKRFFIFF 60  
::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|  
Db 528 SOVIISIIU-LNVAVFVICACIRIYFAVNDELTPA-----NDKYIAKKMAILIF 580

QY 61 TDALCWIP1-FVAKPLSLLOVEIPGTITSWGYGYSAINSLNPLYLTLPFKFE---- 115  
|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|  
Db 591 TDTCPMAP:SFEAISAEKFVLTIIVINSKILLVLYPVNSCANFLYATPKAFQDFLL 640

QY 116 MIHRF 120  
::|::|  
Db 641 LLSRF 645

RESULT 10  
AA9744 lutropin-chorionadotropin receptor precursor - rat  
N:Alternate names: luteinizing hormone-chorionadotropin receptor  
C:Species: Rattus norvegicus (Norway rat)  
C>Date: 03-May-1994 #sequence\_revision 13-Sep-1998 #text\_change 13-Aug-1999  
C:Accession: A49744; A40545; A41343; A61453; A32460  
R:Koo, Y.B.; Ji, I.; Slaughter, R.G.; Ji, T.H.  
Enocrinology 128, 2297-2308, 1991  
A:Title: Structure of the luteinizing hormone receptor gene and multiple exons of the  
A:Reference number: A49744; MUID:91209270; PMID:2019252  
A:Accession: A49744  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-700 <KO>  
A:Cross-references: GB:M68917  
A>Note: authors translated the codon CAA for residue 307 as Glu, AAC for residue 355  
J:Boal-Morris, C.H.; Buczko, E.; Wang, W.; Xie, A.Z.; Durau, M.D.  
J. Biol. Chem. 266, 11355-11359, 1991  
A:Title: Structural organization of the rat luteinizing hormone (LH) receptor gene.  
A:Reference number: A40545; MUID:91250455; PMID:2040640  
A:Accession: A40545  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-320 <SA>  
A:Cross-references: GB:M63918; GB:M63919; GB:M63920; GB:M63921; GB:M63922; GB:M63923;  
R:McFarland, K.C.; Sprengel, R.; Phillips, H.S.; Koehler, M.; Rosenblit, N.; Nikolic  
Science 245, 494-499, 1989  
A:Title: Lutropin-chorionadotropin receptor: an unusual member of the G protein-cou  
A:Reference number: A41343; MUID:89332512; PMID:2502842  
A:Accession: A41343  
A:Molecule type: mRNA  
A:Residues: 1-700 <MF>  
A:Cross-references: GB:M26199; NID:g205178; PIDN:AAA41528.1; RID:g205179  
R:Dufau, M.L.; Minedishi, T.; Buczko, E.S.; Delgado, C.J.; Zhang, R.  
J. Steroid Biochem. 33, 715-720, 1989  
A:Title: Characterization and structure of ovarian and testicular LH/hCG receptors.  
A:Reference number: A61453; MUID:90097014; PMID:2601325  
A:Accession: A61453  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 27-33,'X','35-37,'X','41-44 <DUF>  
R:Roche, P.C.; Ryan, R.J.  
J. Biol. Chem. 264, 4636-4641, 1989  
A:Title: Purification, characterization, and amino-terminal sequence of rat ovarian r  
A:Reference number: A32460; MUID:89147723; PMID:2925659  
A:Accession: A32460  
A:Molecule type: protein  
A:Residues: 27-32,'IX','35-37 <ROC>  
C:Genetics:  
A:Introns: 58/2; 82/2; 107/2; 132/2; 157/2; 183/2; 206/2; 231/2; 293/2; 320/2  
C:Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repea  
C:Keywords: alternative splicing; G protein-coupled receptor; glycoprotein; hormone r  
F;5477/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRRI>

```
; NAME/KEY: SITE
; LOCATION: (133)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-886-52

Query Match          76.5%; Score 632; DB 12; Length 176;
Best Local Similarity 94.0%; Pred. No. 9e-62;
Matches 126; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 AQIYSVAIFLGINLAARFIIIVFSYSGSMFYSHQSATATATEIRNOVKKEMILAKRFFFIYF 60
Db 1 AQIYSVAIFLGINLAARFIIIVFSYSGSMFYSHQSATATATEIRNOVKKEMILAKRFFFIYF 60
QY 61 TDALCWIPFIVAKPLSLQVEIPGTTISWVWIGYSAINSLNPILYLTTRPKEMIHFR 120
Db 61 TDALCWIPFIVAKPLSLQVEIPGTTISWVWIGYSAINSLNPILYLTTRPKEMIHFR 120
QY 121 WHNYRQKSMDSKG 134
Db 121 WHNYRQKSMDSXG 134

RESULT 9
US-10-073-865-84
; Sequence 84, Application US/10073865
; Publication No. US20030044904A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PJ209C1
; TYPE: PRP
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (133)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-073-865-84

Query Match          76.5%; Score 632; DB 15; Length 176;
Best Local Similarity 94.0%; Pred. No. 9e-62;
Matches 126; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 AQIYSVAIFLGINLAARFIIIVFSYSGSMFYSHQSATATATEIRNOVKKEMILAKRFFFIYF 60
Db 1 AQIYSVAIFLGINLAARFIIIVFSYSGSMFYSHQSATATATEIRNOVKKEMILAKRFFFIYF 60
QY 61 TDALCWIPFIVAKPLSLQVEIPGTTISWVWIGYSAINSLNPILYLTTRPKEMIHFR 120
Db 61 TDALCWIPFIVAKPLSLQVEIPGTTISWVWIGYSAINSLNPILYLTTRPKEMIHFR 120
QY 121 WHNYRQKSMDSKG 134
Db 121 WHNYRQKSMDSXG 134

RESULT 10
US-10-103-313-373
; Sequence 373, Application US/10103313
; Publication No. US20030082758A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PJ207C1
; CURRENT APPLICATION NUMBER: US/10/103,313
; CURRENT FILING DATE: 2002-03-12
; NUMBER OF SEQ ID NOS: 653
```

```
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 373
; LENGTH: 176
; TYPE: PRP
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (133)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-103-313-373

Query Match          76.5%; Score 632; DB 15; Length 176;
Best Local Similarity 94.0%; Pred. No. 9e-62;
Matches 126; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 AQIYSVAIFLGINLAARFIIIVFSYSGSMFYSHQSATATATEIRNOVKKEMILAKRFFFIYF 60
Db 1 AQIYSVAIFLGINLAARFIIIVFSYSGSMFYSHQSATATATEIRNOVKKEMILAKRFFFIYF 60
QY 61 TDALCWIPFIVAKPLSLQVEIPGTTISWVWIGYSAINSLNPILYLTTRPKEMIHFR 120
Db 61 TDALCWIPFIVAKPLSLQVEIPGTTISWVWIGYSAINSLNPILYLTTRPKEMIHFR 120
QY 121 WHNYRQKSMDSKG 134
Db 121 WHNYRQKSMDSXG 134

RESULT 11
US-10-073-885-96
; Sequence 96, Application US/10073885
; Publication No. US20030096346A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PJ203C1
; TYPE: PRP
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (133)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-073-885-96

Query Match          76.5%; Score 632; DB 15; Length 176;
Best Local Similarity 94.0%; Pred. No. 9e-62;
Matches 126; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 AQIYSVAIFLGINLAARFIIIVFSYSGSMFYSHQSATATATEIRNOVKKEMILAKRFFFIYF 60
Db 1 AQIYSVAIFLGINLAARFIIIVFSYSGSMFYSHQSATATATEIRNOVKKEMILAKRFFFIYF 60
QY 61 TDALCWIPFIVAKPLSLQVEIPGTTISWVWIGYSAINSLNPILYLTTRPKEMIHFR 120
Db 61 TDALCWIPFIVAKPLSLQVEIPGTTISWVWIGYSAINSLNPILYLTTRPKEMIHFR 120
QY 121 WHNYRQKSMDSKG 134
Db 121 WHNYRQKSMDSXG 134

RESULT 12
US-09-928-175-21
; Sequence 21, Application US/09928175
; Patent No. US20020123618A1
```



```

; GENERAL INFORMATION:
; APPLICANT: Paszty, Christopher J.
; APPLICANT: Gong, Jianhua
; APPLICANT: Daugherty, Betsy
; APPLICANT: Rogers, No. US20020123618Alma
; TITLE OF INVENTION: Leucine-Rich G Protein Coupled Receptor-8 Molecules and
; FILE REFERENCE: 00-1229
; CURRENT APPLICATION NUMBER: US/09/928,175
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/224,455
; PRIOR FILING DATE: 2000-08-10
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 21
; LENGTH: 718
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-928-175-21

Query Match          49.5%; Score 409; DB 10; Length 718;
Best Local Similarity 58.9%; Pred. No. 2e-36;
Matches 73; Conservative 27; Mismatches 24; Indels 0; Gaps 0;

QY 4 YSVAIFLGINLAFTIIVFSYSGMSFYSVHQSATATATIRNOVKKEMILAKRFFFIIVFTDA 63
DB 553 YSLGIFLVNLLAFIVISYVTMFCSTHKTALQTAQVRSHGKEVAVANRFFFIIVFSDA 612

QY 64 LCWPIFVAKPLSLLOVEIPGTITSWVIGYSAINSLNPILYTLTTRPKEMIHFRFHN 123
DB 613 ICWIPFVVKILSLLOVEIPGTITSWIVFFLPVNSALNPILYTLTTSFFKDKLKOLLHK 672

QY 124 YRQR 127
DB 673 HRRK 676

RESULT 13
US-09-928-175-20
; Sequence 20, Application US/09928175
; Patent No. US20020123618Al
; GENERAL INFORMATION:
; APPLICANT: Paszty, Christopher J.
; APPLICANT: Gong, Jianhua
; APPLICANT: Daugherty, Betsy
; APPLICANT: Rogers, No. US20020123618Alma
; TITLE OF INVENTION: Leucine-Rich G Protein Coupled Receptor-8 Molecules and
; FILE REFERENCE: 00-1229
; CURRENT APPLICATION NUMBER: US/09/928,175
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/224,455
; PRIOR FILING DATE: 2000-08-10
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 737
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-928-175-20

Query Match          49.5%; Score 409; DB 10; Length 737;
Best Local Similarity 58.9%; Pred. No. 2e-36;
Matches 73; Conservative 27; Mismatches 24; Indels 0; Gaps 0;

QY 4 YSVAIFLGINLAFTIIVFSYSGMSFYSVHQSATATATIRNOVKKEMILAKRFFFIIVFTDA 63
DB 572 YSLGIFLVNLLAFIVISYVTMFCSTHKTALQTAQVRSHGKEVAVANRFFFIIVFSDA 631

QY 64 LCWPIFVAKPLSLLOVEIPGTITSWVIGYSAINSLNPILYTLTTRPKEMIHFRFHN 123
DB 632 ICWIPFVVKILSLLOVEIPGTITSWIVFFLPVNSALNPILYTLTTSFFKDKLKOLLHK 691

```

```

QY 124 YRQR 127
DB 692 HRRK 695

RESULT 14
US-10-229-735-3
; Sequence 3, Application US/10229735
; Publication No. US20030082650A1
; GENERAL INFORMATION:
; APPLICANT: Baylor College of Medicine
; APPLICANT: Agoulnik, Alexander I.
; TITLE OF INVENTION: The Great Gene and Protein
; FILE REFERENCE: 7572/73263
; CURRENT APPLICATION NUMBER: US/10/229,735
; CURRENT FILING DATE: 2002-08-29
; PRIOR APPLICATION NUMBER: 60/315,696
; PRIOR FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: 60/351,432
; PRIOR FILING DATE: 2002-01-28
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 737
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-229-735-3

Query Match          49.5%; Score 409; DB 15; Length 737;
Best Local Similarity 58.9%; Pred. No. 2e-36;
Matches 73; Conservative 27; Mismatches 24; Indels 0; Gaps 0;

QY 4 YSVAIFLGINLAFTIIVFSYSGMSFYSVHQSATATATIRNOVKKEMILAKRFFFIIVFTDA 63
DB 572 YSLGIFLVNLLAFIVISYVTMFCSTHKTALQTAQVRSHGKEVAVANRFFFIIVFSDA 631

QY 64 LCWPIFVAKPLSLLOVEIPGTITSWVIGYSAINSLNPILYTLTTRPKEMIHFRFHN 123
DB 632 ICWIPFVVKILSLLOVEIPGTITSWIVFFLPVNSALNPILYTLTTSFFKDKLKOLLHK 691

QY 124 YRQR 127
DB 692 HRRK 695

RESULT 15
US-10-321-807-18
; Sequence 18, Application US/10321807
; Publication No. US20030166148A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Rupong
; APPLICANT: Dang, Ruong T.
; APPLICANT: Lowitz, Kevin P.
; TITLE OF INVENTION: No. US20030166148A1-Endogenous, Constitutively Activated Human
; TITLE OF INVENTION: Receptors
; FILE REFERENCE: AREN0086
; CURRENT APPLICATION NUMBER: US/10/321,807
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: US/09/714,008
; PRIOR FILING DATE: 2000-11-16
; PRIOR APPLICATION NUMBER: 60/166,088
; PRIOR FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: PCT/US99/23938
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: 60/166,088
; PRIOR FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: 60/166,099
; PRIOR FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: 60/166,369
; PRIOR FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: 60/171,902
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: 60/171,901

```

```
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: 60/171,900
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: 60/181,749
; PRIOR FILING DATE: 2000-02-11
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 133
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 18
; LENGTH: 355
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-321-807-18

Query Match      47.5%; Score 392.5; DB 12; Length 355;
Best Local Similarity 59.1%; Pred. No. 5.4e-35;
Matches 75; Conservative 25; Mismatches 26; Indels 1; Gaps 1;

QY      4 YVAIFLGINLAAFIIIVFSYSGMFYVHQSAITATEIRNOVKKEMILAKRFFFIIVTDA 63
Db      190 YSLGIFLGYNLIATLIIVFSYITMFCISIOKTALQTTEVRNCFGREVAVANRFFIIVSDA 249

QY      64 LCHWPIFYAKPLSLLOVEIPGTITTSWVWVGYSAINSALNPILYITLITRPFKEMIHFWHN 123
Db      250 ICWIPFVVKILSLFRVEIPDTMTSWIVIFLFLVNSALNPILYITLITNPFKDLKQLLHK 309

QY      124 YRQKSM 130
Db      310 H-QRKSI 315
```

Search completed: September 25, 2003, 15:10:13  
Job time : 65 secs

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 25, 2003, 14:56:36 : Search time 30 Seconds  
(without alignments)  
221.427 Million cell updates/sec

Title: US-10-049-568-2  
Perfect score: 826  
Sequence: 1 AQIYSVAIFLGINLAFFII.....HMLHSSGGKGGCRRCHLS 157

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: Issued Patents: AA:\*

1: /cn2.6/ptodata/1/iaa/5A\_COMB.pep.\*  
2: /cn2.6/ptodata/1/iaa/5B\_COMB.pep.\*  
3: /cn2.6/ptodata/1/iaa/6A\_COMB.pep.\*  
4: /cn2.6/ptodata/1/iaa/6B\_COMB.pep.\*  
5: /cn2.6/ptodata/1/iaa/PCUS\_COMB.pep.\*  
6: /cn2.6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	147.5	17.9	792	3	US-07-741-453A-56
2	140.5	17.0	696	3	US-07-757-342D-4
3	139.5	16.9	692	3	US-07-757-342D-6
4	137.5	16.6	795	3	US-07-741-453A-55
5	136	16.5	370	1	US-08-118-270-18
6	136	16.5	370	5	PCT-US93-08528-18
7	134	16.2	700	3	US-07-757-342D-3
8	131	15.9	611	3	US-07-757-342D-8
9	131	15.9	636	3	US-07-757-342D-7
10	131	15.9	674	3	US-07-757-342D-10
11	131	15.9	699	3	US-07-757-342D-2
12	125.5	15.2	420	4	US-08-795-876-33
13	125.5	15.2	423	4	US-08-795-876-38
14	125.5	15.2	436	4	US-08-795-876-2
15	125.5	15.2	695	1	US-08-487-886-2
16	125.5	15.2	695	3	US-08-482-855-2
17	125.5	15.2	695	4	US-08-474-986-2
18	124	15.0	338	3	US-08-988-876-8
19	124	15.0	338	3	US-09-303-524A-2
20	123.5	15.0	325	2	US-08-467-948A-29
21	123.5	15.0	325	3	US-08-467-947A-29
22	122	14.8	764	3	US-07-741-453A-54
23	120.5	14.6	764	3	US-07-741-453A-60
24	119.5	14.5	332	1	US-08-293-563-5
25	119.5	14.5	378	3	US-09-082-088-2
26	119.5	14.5	378	4	US-09-346-117-2
27	119.5	14.5	378	4	US-09-169-205D-22

28	119.5	14.5	378	4	US-09-731-030A-18
29	119.5	14.5	517	2	US-08-467-568-10
30	119.5	14.5	517	2	US-09-030-582-10
31	118.5	14.3	515	1	US-08-194-338-3
32	118	14.3	764	3	US-07-741-453A-61
33	117.5	14.2	336	1	US-08-118-370-54
34	117.5	14.2	336	5	PCT-US93-08528-54
35	117.5	14.2	420	3	US-09-255-368-6
36	117.5	14.2	513	2	US-08-406-855A-21
37	117.5	14.2	513	3	US-09-206-899-21
38	117.5	14.2	515	1	US-08-444-734A-7
39	117.5	14.2	515	1	US-08-722-001-25
40	117.5	14.2	515	2	US-08-406-855A-22
41	117.5	14.2	515	3	US-09-206-899-22
42	117.5	14.2	515	3	US-09-032-742-5
43	117.5	14.2	515	4	US-09-688-415-9
44	117.5	14.2	515	4	US-09-688-415-10
45	117.5	14.2	764	3	US-07-757-342D-5

## ALIGNMENTS

RESULT 1  
US-07-741-453A-56  
; Sequence 56, Application US/07741453A  
; Patent No. 6228597  
; GENERAL INFORMATION:  
; APPLICANT: PARMENTIER, MARC  
; APPLICANT: LIBERT, FREDERIC  
; APPLICANT: DUMONT, JACQUES  
; APPLICANT: VASSART, GILBERT  
; TITLE OF INVENTION: POLYPEPTIDES HAVING THYROTROPIN-RECEPTOR  
; TITLE OF INVENTION: ACTIVITY, NUCLEIC ACID SEQUENCES CODING FOR SUCH RECEPTORS  
; TITLE OF INVENTION: AND POLYPEPTIDES, AND APPLICATIONS OF THESE POLYPEPTIDES  
; NUMBER OF SEQUENCES: 62  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: CUSHMAN DABY & CUSHMAN  
; STREET: 1615 L STREET, N.W.  
; CITY: WASHINGTON, D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07741453A  
; FILING DATE: 19911015  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: KOKOLIS, PAUL N.  
; REGISTRATION NUMBER: 16773  
; REFERENCE/DOCKET NUMBER: 91913/1107/US/ST  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 861-3000  
; TELEFAX: (202) 822-0944  
; TELEX: 6714827 CUSH  
; INFORMATION FOR SEQ ID NO: 56:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 792 amino acids  
; TYPE: AMINO ACID  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-07-741-453A-56

Query Match 17.9%; Score 147.5; DB 3; Length 792;  
Best Local Similarity 25.9%; Pred. No. 1.1e-07;  
Matches 42; Conservative 27; Mismatches 60; Indels 33; Gaps 4;  
1 AQIYSVAIFL----GINLAFFIIIVSGMFYSVHQSAITATEIRNOYKEMILAKRF 55

```

Db 596 SQVYLITLIPICGCLNVAFIICACIKIYFVQNPFLMAT-----NKDTAKAKM 649
QY 56 FPIVETDLCWIPFVAKPLSLQVE-----IPGTITSWVIGYSAINSALNPILYTL 108
Db 650 AVLIPTFTCMATIGCHGCLSFALSALKVPLITVTNSKVLILVFPVNSCANPFYAI 709
QY 109 TTRPFKEMIRFWHNYRQRKMSDKGIRKMLHSSGCKGH 150
Db 710 FTKAFFR-----DPICGCGFFLLSKSGCKH 736

```

## RESULT 2

```

US-07-757-342D-4
; Sequence 4, Application US/07757342D
; Patent No. 6218509
; GENERAL INFORMATION:
; APPLICANT: IGARASHI, Masao
; MINEGISHI, Takashi
; NAKAMURA, Kazuto
; TITLE OF INVENTION: PROTEIN, DNA AND USE THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &
; CUSHMAN
; STREET: 130 Water Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: US
; ZIP: 02109

```

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/757,342D
FILING DATE: 10-Sep-1991
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: BUCKLEY, Linda M.
REGISTRATION NUMBER: 31003
REFERENCE/DOCKET NUMBER: 41226
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)523-3400
TELEFAX: (617)523-6440
TELEX: 200291 STRE UR

```

```

INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 696 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 4:

```

```

US-07-757-342D-4
Query Match 17.0%; Score 140.5; DB 3; Length 696;
Best Local Similarity 27.2%; Pred. No. 5.6e-07;
Matches 41; Conservative 26; Mismatches 55; Indels 29; Gaps 4;

```

```

QY 1 AQIYSAVLFGINLAFLIIVSYGSMFYSHQSAITATEIRNQVKEMILAKRFFFI 60
Db 524 SQVYLITLILNVAFIICACIKIYFVQNPFLMAT-----NKDTAKAKMVLIF 576
QY 61 TDLACWIPFVAKPLSLQVEIPGTITSWVIGYSAINSALNPILYTLTRPFKE 119
Db 577 TDFTCMAPISFALSALKVPLITVTNSKVLILVFPVNSCANPFYAIFTKAER 631
QY 120 FWHNYRQRKMSDKGIRKMLHSSGCKGH 150
Db 632 -----RDFLLSKSGCKH 646

```

```

Query Match 17.0%; Score 140.5; DB 3; Length 696;
Best Local Similarity 27.2%; Pred. No. 5.6e-07;
Matches 41; Conservative 26; Mismatches 55; Indels 29; Gaps 4;

```

```

QY 1 AQIYSAVLFGINLAFLIIVSYGSMFYSHQSAITATEIRNQVKEMILAKRFFFI 60
Db 524 SQVYLITLILNVAFIICACIKIYFVQNPFLMAT-----NKDTAKAKMVLIF 576
QY 61 TDLACWIPFVAKPLSLQVEIPGTITSWVIGYSAINSALNPILYTLTRPFKE 119
Db 577 TDFTCMAPISFALSALKVPLITVTNSKVLILVFPVNSCANPFYAIFTKAER 631
QY 120 FWHNYRQRKMSDKGIRKMLHSSGCKGH 150
Db 632 -----RDFLLSKSGCKH 646

```

```

RESULT 3
US-07-757-342D-6
; Sequence 6, Application US/07757342D
; Patent No. 6218509
; GENERAL INFORMATION:
; APPLICANT: IGARASHI, Masao
; MINEGISHI, Takashi
; NAKAMURA, Kazuto
; TITLE OF INVENTION: PROTEIN, DNA AND USE THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &
; CUSHMAN
; STREET: 130 Water Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: US
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/757,342D
; FILING DATE: 10-Sep-1991
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: BUCKLEY, Linda M.
; REGISTRATION NUMBER: 31003
; REFERENCE/DOCKET NUMBER: 41226
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)523-3400
; TELEFAX: (617)523-6440
; TELEX: 200291 STRE UR
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 692 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
;
US-07-757-342D-6
Query Match 16.9%; Score 139.5; DB 3; Length 692;
Best Local Similarity 27.3%; Pred. No. 7.1e-07;
Matches 47; Conservative 30; Mismatches 60; Indels 35; Gaps 6;

```

```

QY 1 AQIYSAVLFGINLAFLIIVSYGSMFYSHQSAITATEIRNQVKEMILAKRFFFI 60
Db 526 SQIYVMAL-LVNLVAFVVICGTYHITVVRNPITVSS-----SSDTKAKRMALIF 578
QY 61 TDLACWIPFVAKPLSLQVEIPGTITSWVIGYSAINSALNPILYTLTRPFKE 115
Db 579 TDFLCMAPISFALSALKVPLITVTNSKVLILVFPVNSCANPFYAIFTKNFRDFFI 638
QY 116 MLHFR-----WHNYRQRKMSDKGIRK-----MLHSS 144
Db 639 LLKFGCYEQAIQYRTSTSTATHNFARKSHCSAPPRTVNSYVLVPLNHSS 690

```

## RESULT 4

```

US-07-741-453A-55
; Sequence 55, Application US/07741453A
; Patent No. 6228597
; GENERAL INFORMATION:

```

```

; APPLICANT: PARMENTIER, MARC
; APPLICANT: LIBERT, FREDERIC
; APPLICANT: DOMONT, JACQUES
; APPLICANT: VASSART, GILBERT
; TITLE OF INVENTION: POLYPEPTIDES HAVING THYROTROPIN-RECEPTOR
; ACTIVITY, NUCLEIC ACID SEQUENCES CODING FOR SUCH RECEPTORS
; AND POLYPEPTIDES, AND APPLICATIONS OF THESE POLYPEPTIDES

```

QY	1	AQIYSAIFGIGNLAFAIIIVFSYGSMEYSVHQSAITATEIRNQVKEMILAKREFFIVF	60
Db	1	AQIYSAIFGIGNLAFAIIIVFSYGSMEYSVHQSAITATEIRNQVKEMILAKREFFIVF	60
QY	61	TDALCWPFIYVAKPLSLLQVEIPGTITSWVVIYGYSAINSALNPILYLTITRPFKEMIHFR	120
Db	61	TDALCWPFIYVAKPLSLLQVEIPGTITSWVVIYGYSAINSALNPILYLTITRPFKEMIHFR	120
QY	121	WHNYRQRKSMDSKG	134
		!::	
		!::	
Db	121	WHNYRQRKSMDSKG	134

## RESULTS

```

US-09-764-877-2009
; Sequence 2009, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764,877
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2009
; LENGTH: 176
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (133)
; OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids
US-09-764-877-2009

```

Query Match 76.5%; Score 632; DB 10; Length 176;  
Best Local Similarity 94.0%; Pred. No. 9e-62;  
Matches 126; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY	1	AQIYVAFLGINLAFLIIIVSGSMFYSHVOSAITATATETNRQVKEMILAKREFPIVF	60
Db	1	AQIYVAFLGINLAFLIIIVSGSMFYSHVOSAITATATETNRQVKEMILAKREFPIVF	60
QY	61	TDALCWIPFVAKPLSLQLOVEIPGTTISWVVGYSAINALNPILYTLTRPFKEMIHRE	120
Db	61	TDALCWIPFVVKPLSLQLOVEIPGTTISWVVFILPINSALNPILYTLTRPFKEMIHRE	120
QY	121	WHNYRQKSMDSXG	134
		!:::	
Db	121	WYNYRQKSMDSXG	134
		!:::	

## RESULT 7

```

US-09-989-442-105
Sequence 105, Application US/09989442
Publication No. US20030013649A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: P0208
CURRENT APPLICATION NUMBER: US/09/989,442
CURRENT FILING DATE: 2001-11-21
PRIORITY APPLICATION NUMBER: 60/179,065
PRIORITY FILING DATE: 2000-01-31
PRIORITY APPLICATION NUMBER: 60/180,628
PRIORITY FILING DATE: 2000-02-04
PRIORITY APPLICATION NUMBER: 60/214,886
PRIORITY FILING DATE: 2000-06-28
PRIORITY APPLICATION NUMBER: 60/217,487
PRIORITY FILING DATE: 2000-07-11
PRIORITY APPLICATION NUMBER: 60/225,758
PRIORITY FILING DATE: 2000-09-05
PRIORITY APPLICATION NUMBER: 60/251,868
PRIORITY FILING DATE: 2000-12-08
PRIORITY APPLICATION NUMBER: 60/229,344
PRIORITY FILING DATE: 2000-09-01
PRIORITY APPLICATION NUMBER: 60/234,997
PRIORITY FILING DATE: 2000-09-25
PRIORITY APPLICATION NUMBER: 60/229,343
PRIORITY FILING DATE: 2000-09-01
PRIORITY APPLICATION NUMBER: 60/229,345
PRIORITY FILING DATE: 2000-09-01
PRIORITY APPLICATION NUMBER: 60/229,287
PRIORITY FILING DATE: 2000-09-01
PRIORITY APPLICATION NUMBER: 60/229,513
PRIORITY FILING DATE: 2000-09-05
PRIORITY APPLICATION NUMBER: 60/231,413
PRIORITY FILING DATE: 2000-09-08
PRIORITY APPLICATION NUMBER: 60/229,509
PRIORITY FILING DATE: 2000-09-05

```

; PRIOR APPLICATION NUMBER: 60/236,367  
; PRIOR FILING DATE: 2000-09-29  
; PRIOR APPLICATION NUMBER: 60/237,039  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR APPLICATION NUMBER: 60/237,038  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR APPLICATION NUMBER: 60/236,370  
; PRIOR FILING DATE: 2000-09-29  
; PRIOR APPLICATION NUMBER: 60/236,802  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR APPLICATION NUMBER: 60/237,037  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR APPLICATION NUMBER: 60/237,040  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR APPLICATION NUMBER: 60/240,960  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/239,935  
; PRIOR FILING DATE: 2000-10-13  
; PRIOR APPLICATION NUMBER: 60/239,937  
; PRIOR FILING DATE: 2000-10-13  
; PRIOR APPLICATION NUMBER: 60/241,787  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/246,474  
; PRIOR FILING DATE: 2000-11-08  
; PRIOR APPLICATION NUMBER: 60/246,532  
; PRIOR FILING DATE: 2000-11-08  
; PRIOR APPLICATION NUMBER: 60/249,216  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,210  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/226,681  
; PRIOR FILING DATE: 2000-08-22  
; PRIOR APPLICATION NUMBER: 60/225,759  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/225,213  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/227,182  
; PRIOR FILING DATE: 2000-08-22  
; PRIOR APPLICATION NUMBER: 60/225,214  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/235,836  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/230,438  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/215,135  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: 60/225,266  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/249,218  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,208  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,213  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,212  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,207  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,245  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,244  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,217  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,211  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,215  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,264  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,214  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,297

; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/232,400  
; PRIOR FILING DATE: 2000-09-14  
; PRIOR APPLICATION NUMBER: 60/231,242  
; PRIOR FILING DATE: 2000-09-08  
; PRIOR APPLICATION NUMBER: 60/232,081  
; PRIOR FILING DATE: 2000-09-08  
; PRIOR APPLICATION NUMBER: 60/232,080  
; PRIOR FILING DATE: 2000-09-08  
; PRIOR APPLICATION NUMBER: 60/231,414  
; PRIOR FILING DATE: 2000-09-08  
; PRIOR APPLICATION NUMBER: 60/231,244  
; PRIOR FILING DATE: 2000-09-08  
; PRIOR APPLICATION NUMBER: 60/233,064  
; PRIOR FILING DATE: 2000-09-14  
; PRIOR APPLICATION NUMBER: 60/233,063  
; PRIOR FILING DATE: 2000-09-14  
; PRIOR APPLICATION NUMBER: 60/232,397  
; PRIOR FILING DATE: 2000-09-14  
; PRIOR APPLICATION NUMBER: 60/232,399  
; PRIOR FILING DATE: 2000-09-14  
; PRIOR APPLICATION NUMBER: 60/232,401  
; PRIOR FILING DATE: 2000-09-14  
; PRIOR APPLICATION NUMBER: 60/241,808  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/241,826  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/241,786  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/241,221  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/246,475  
; PRIOR FILING DATE: 2000-11-08  
; PRIOR APPLICATION NUMBER: 60/231,243  
; PRIOR FILING DATE: 2000-09-08  
; PRIOR APPLICATION NUMBER: 60/233,065  
; PRIOR FILING DATE: 2000-09-14

Query Match 76.5%; Score 632; DB 11; Length 176;  
Best Local Similarity 94.0%; Pred. No. 9e-62;  
Matches 126; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

Qy 1 ACISVAIFLGINLAFLIIVFSYSGMSYVHQSATTAIRNOVKKEMILAKRFFIVF 60  
Db 1 ACISVAIFLGINLAFLIIVFSYSGMSYVHQSATTAIRNOVKKEMILAKRFFIVF 60  
Qy 61 TDALCWIPIFVAKPLSLQVEIPGTTTSVVVIGYSAINSPILYTLTTRPFKEMIHFR 120  
Db 61 TDALCWIPIFVAKPLSLQVEIPGTTTSVVVIFILPINSALNPILYTLTTRPFKEMIHFR 120  
Qy 121 WHNYRQRKSMDSKG 134  
Db 121 WHNYRQRKSMDSKG 134

RESULT 8  
US-09-764-886-52  
; Sequence 52, Application US/09764886  
; Publication NO. US20030139327A9  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: Ptz02  
; CURRENT APPLICATION NUMBER: US/09/764,886  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - consult PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 88  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 52  
; LENGTH: 176  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:

```

CITY: Boston
STATE: Massachusetts
COUNTRY: US
ZIP: 02109

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/757,342D
FILING DATE: 10-Sep-1991
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: BUCKLEY, Linda M.
REGISTRATION NUMBER: 31003
REFERENCE/DOCKET NUMBER: 41226
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)523-3400
TELEFAX: (617)523-6440
TELEX: 200291 STRE UR

INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 674 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-07-757-342D-10

Query Match      15.9%; Score 131; DB 3; Length 674;
Best Local Similarity 29.6%; Pred. No. 5.8e-06;
Matches 34; Conservative 25; Mismatches 49; Indels 8; Gaps 3

QY    1 AQIYSVAIFGINLARFIILVFSGMFYSVHCSAITATEIRNQVKEMILAKREFEIVF 60
Db    :|: |: |: |: || |: |||| |: |||| |: |||| |: |||| |: |||| |: ||||
       499 SQWYLITLI-LNWAVFFICACIKIFYVRNPELMAT-----NRDKPKIAKRAMLII 551

QY    61 TDLACWIPI-FVAKPLSLQLVEIPGTITSWWVGYSAINSNALNPILYTLTTRPFK 114
Db    == || || || || || || || || || || || || || || || || || || || ||
       552 TDFTCMARISFEALSAAKRVPLITVTNSKVLLVLPINSANCFLYAIFTTFEQ 606

RESULT 11
US-07-757-342D-2
Sequence 2, Application US/07757342D
Patent No. 6218509
GENERAL INFORMATION:
APPLICANT: IGARASHI, Masao
MINEGISHI, Takashi
NAKANURA, Kazuto
TITLE OF INVENTION: PROTEIN, DNA AND USE THEREOF
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESSES:
ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS & CUSHMAN
STREET: 130 Water Street
CITY: Boston
STATE: Massachusetts
COUNTRY: US
ZIP: 02109

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/757,342D
FILING DATE: 10-Sep-1991
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: BUCKLEY, Linda M.
REGISTRATION NUMBER: 31003

```

REFERENCE/DOCKET NUMBER: 41226  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)523-3400  
TELEFAX: (617)523-6440  
TELEX: 200291 STRE UR  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 699 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-07-757-342D-2

Query Match 15.9%; Score 131; DB 3; Length 699;  
Best Local Similarity 29.6%; Pred. No. 6e-06; Indels 8; Gaps 3;  
Matches 34; Conservative 25; Mismatches 48;  
QY 1 AQIYSVAIFGLINLAAPFIIIVFSGMFSYVHQSATATATIRNOVKKEMILAKRFFFI 60  
DB 524 SQVYLITLIL-LNVVAFIICACYIKIYFAVENPELMAT-----NKDTRIAKRAMLIF 576  
QY 61 TDALCWIPIT-FVAKPLSLQVEIPGTTTSWVIGYSAINSLNPILYTLTTRPFK 114  
DB 577 TDTCMAPISFAISAKVPLITVTSKVLVLFHPINSCANPFLYAITKFRDFFI 631

RESULT 12  
US-08-795-876-33  
Sequence 33, Application US/08795876  
Patent No. 6403305  
GENERAL INFORMATION:  
APPLICANT: Gershengorn, Marvin C.  
APPLICANT: Geras-Raaka, Elizabeth  
APPLICANT: Nussenzeig, Daniel R.  
TITLE OF INVENTION: STRATEGY TO CLONE DRUGS FOR G PROTEIN  
TITLE OF INVENTION: COUPLED RECEPTORS  
NUMBER OF SEQUENCES: 44  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: NIXON, HARGRAVE, DEVANS & DOYLE LLP  
STREET: Clinton Square, P.O. Box 1051  
CITY: Rochester  
STATE: New York  
COUNTRY: USA  
ZIP: 14603  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/795,876  
FILING DATE:  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: BRAMAN, SUSAN J.  
REGISTRATION NUMBER: 34,103  
REFERENCE/DOCKET NUMBER: 19603/1280  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 716-263-1636  
TELEFAX: 716-263-1600  
INFORMATION FOR SEQ ID NO: 33:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 420 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-795-876-33

Query Match 15.2%; Score 125.5; DB 4; Length 420;  
Best Local Similarity 26.1%; Pred. No. 1.3e-05;  
Matches 43; Conservative 30; Mismatches 59; Indels 33; Gaps 6;

QY 1 AQIYSVAIFGLINLAAPFIIIVFSGMFSYVHQSATATATIRNOVKKEMILAKRFFFI 60  
DB 252 SQIYVMSL-LVLNVLAFWICGCIHIYLTVRNPNISS-----SSDTRIAKRAMLIF 304  
QY 61 TDALCWIPIT-FVAKPLSLQVEIPGTTTSWVIGYSAINSLNPILYTLTTRPFK----- 114  
DB 305 TDFLCMAPISFAISAKVPLITVTSKVLVLFHPINSCANPFLYAITKFRDFFI 364  
QY 115 -----EMHFRFWHNYRQKSMDSKIRKHMHLHSSGGKCGHC 151  
DB 365 LLSKCGCYEMAQI---YRTETSTV-----HNTHPRNGHC 397  
RESULT 13  
US-08-795-876-38  
Sequence 38, Application US/08795876  
Patent No. 6403305  
GENERAL INFORMATION:  
APPLICANT: Gershengorn, Marvin C.  
APPLICANT: Geras-Raaka, Elizabeth  
APPLICANT: Nussenzeig, Daniel R.  
TITLE OF INVENTION: STRATEGY TO CLONE DRUGS FOR G PROTEIN  
TITLE OF INVENTION: COUPLED RECEPTORS  
NUMBER OF SEQUENCES: 44  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: NIXON, HARGRAVE, DEVANS & DOYLE LLP  
STREET: Clinton Square, P.O. Box 1051  
CITY: Rochester  
STATE: New York  
COUNTRY: USA  
ZIP: 14603  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/795,876  
FILING DATE:  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: BRAMAN, SUSAN J.  
REGISTRATION NUMBER: 34,103  
REFERENCE/DOCKET NUMBER: 19603/1280  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 716-263-1636  
TELEFAX: 716-263-1600  
INFORMATION FOR SEQ ID NO: 38:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 423 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-795-876-38

Query Match 15.2%; Score 125.5; DB 4; Length 423;  
Best Local Similarity 26.1%; Pred. No. 1.3e-05;  
Matches 43; Conservative 30; Mismatches 59; Indels 33; Gaps 6;  
QY 1 AQIYSVAIFGLINLAAPFIIIVFSGMFSYVHQSATATATIRNOVKKEMILAKRFFFI 60  
DB 255 SQIYVMSL-LVLNVLAFWICGCIHIYLTVRNPNISS-----SSDTRIAKRAMLIF 307  
QY 61 TDALCWIPIT-FVAKPLSLQVEIPGTTTSWVIGYSAINSLNPILYTLTTRPFK----- 114  
DB 308 TDFLCMAPISFAISAKVPLITVTSKVLVLFHPINSCANPFLYAITKFRDFFI 367  
QY 115 -----EMHFRFWHNYRQKSMDSKIRKHMHLHSSGGKCGHC 151  
DB 368 LLSKCGCYEMAQI---YRTETSTV-----HNTHPRNGHC 400



QY 1 AQYISVAIFLGINLAAPILIVFSYSGMFYSVHQSAITATEIRNOVKKEMILAKREFFIVF 60  
Db 524 SQVITITIL-LNVVAFIICACYIXIYFAVRNPELMAT-----NKDTKIARKMAILIF 576  
QY 61 TDALCWIPI-FVAKPLSILQVEIPGWTITSVVIGYSAINALNPILYTLTTRPEK 114  
Db 577 TDFTCMAPISFFAISAAKVPPLITVTNSKVLVLVEYPINSCANPFYIAITFTFQ 631

Search completed: September 25, 2003, 15:01:39  
Job time : 20 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 25, 2003, 14:55:26 ; Search time 36 Seconds  
(without alignments)  
1125.396 Million cell updates/sec

Title: US-10-049-568-2  
Perfect score: 826  
Sequence: 1 AQTYSVAIFGLINLAFFII.....HMLHSSGGKCGHRRCHLS 157

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SPTREMBL\_23:\*

- 1: sp.archaea:\*
- 2: sp.bacteria:\*
- 3: sp.fungi:\*
- 4: sp.human:\*
- 5: sp.invertebrate:\*
- 6: sp.mammal:\*
- 7: sp.mnc:\*
- 8: sp.organelle:\*
- 9: sp.phage:\*
- 10: sp.plant:\*
- 11: sp.rodent:\*
- 12: sp.virus:\*
- 13: sp.vertebrate:\*
- 14: sp.unclassified:\*
- 15: sp.virus:\*
- 16: sp.bacteriap:\*
- 17: sp.archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query %	Match	Length	DB	ID	Description
1	224.5	27.2	334	5	Q9VVG0		Q9vvg0 drosophila
2	212.5	25.7	765	5	Q9VBP0		Q9vbp0 drosophila
3	154	18.6	693	13	Q9DGC5		Q9dgc5 oreochromis
4	148	17.9	1012	5	Q95YI6		Q95yi6 asterina pe
5	148	17.9	1280	5	Q95YI7		Q95yi7 asterina pe
6	143.5	17.4	696	13	Q9GT85		Q9gt85 ictalurus p
7	139.5	16.9	688	11	Q64183		Q64183 rattus sp.
8	138	16.7	307	13	Q90XC7		Q90xc7 salmo salar
9	134	16.2	814	13	Q91949		Q91949 oncorhynch
10	133.5	16.2	424	13	Q9DPE3		Q9dpe3 rana catesb
11	132	16.0	701	6	Q8SPS8		Q8sps8 bos taurus
12	132	16.0	793	13	Q91948		Q91948 oncorhynch
13	131	15.9	779	13	Q918N7		Q918n7 morone saxa
14	130.5	15.8	778	13	Q98TF4		Q98tf4 oreochromis
15	130	15.7	1050	5	Q9BN18		Q9bn18 drosophila
16	130	15.7	1300	5	Q9NKD6		Q9nkD6 drosophila

17	130	15.7	1360	5	Q9NDI1		Q9ndi1 drosophila
18	128	15.5	658	13	Q9PVN9		Q9pvn9 oncorhynch
19	127	15.4	662	13	Q98T84		Q98t84 ictalurus p
20	125.5	15.2	410	4	Q16225		Q16225 homo sapien
21	125	15.1	662	13	Q9PW16		Q9pw16 clarias gar
22	124	15.0	338	4	Q81YT7		Q81yt7 homo sapien
23	124	15.0	696	13	Q9DGF5		Q9dgf5 cynops pyrr
24	123.5	15.0	695	11	Q8R428		Q8r428 cavia porce
25	123	14.9	335	6	Q46387		Q46387 mustela vis
26	123	14.9	710	13	Q8AXM5		Q8axm5 clarias gar
27	122.5	14.8	364	5	Q819M7		Q819m7 asterina mi
28	122	14.8	793	6	Q95MF7		Q95mf7 canis famil
29	122	14.8	763	6	Q9BGN4		Q9bgn4 felis silve
30	121	14.6	601	13	Q42500		Q42500 meleagris g
31	121	14.6	764	11	Q9D697		Q9d697 mus musculu
32	119.5	14.5	384	13	Q9PUQ8		Q9puq8 figu rubrip
33	119.5	14.5	398	4	Q9H228		Q9h228 homo sapien
34	119.5	14.5	459	11	Q8CG15		Q8cg15 mus musculu
35	119.5	14.5	515	11	Q9DBL0		Q9dbl0 mus musculu
36	119.5	14.5	724	13	Q9PVP0		Q9pvp0 oncorhynch
37	119	14.4	739	6	Q9BG55		Q9bg55 sus scrofa
38	119	14.4	764	6	Q8SP9		Q8sp9 sus scrofa
39	119	14.4	764	6	Q9BG56		Q9bg56 sus scrofa
40	117.5	14.2	418	13	Q90WJ2		Q90wj2 xenopus lae
41	117.5	14.2	470	4	O43200		O43200 homo sapien
42	117.5	14.2	533	4	Q9PIV4		Q9piv4 homo sapien
43	117.5	14.2	764	6	O8HY54		O8hy54 cercopithec
44	117.5	14.2	764	6	O8HY53		O8hy53 macaca mula
45	116	14.0	296	11	Q8BK41		Q8br41 mus musculu

## ALIGNMENTS

RESULT 1

Q9VVG0 PRELIMINARY; PRT; 334 AA.

ID Q9VVG0

AC Q9VVG0;

DT 01-MAY-2000 (TREMREL. 13, Created)

DT 01-MAY-2000 (TREMREL. 13, Last sequence update)

DT 01-MAR-2003 (TREMREL. 23, Last annotation update)

DE CG4187 protein.

GN CG4187.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

ON NCBI\_TaxID=7227;

RX MEDLINE=20196006; PubMed=10731132;

RC STRAIN=BERKELEY;

RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,

RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,

RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,

RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,

RA Abril J.F., Agayani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,

RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,

RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,

RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,

RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,

RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,

RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,

RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,

RA Fostier C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,

RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,

RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,

RA Hostin D., Houston K.A., Howland T.J., Wei W.-H., Ibegwam C.,

RA Jalali M., Kalush F., Karpén G.H., Ke Z., Kennison J.A., Ketchum K.A.,

RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
Dodson K., Doup L.E., Downes N., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
Durbini K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,  
Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,  
Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,  
Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
Kimmel B.E., Kodira C.C., Kraft C., Kravitz S., Rulp D., Lai Z.,  
Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
Merkulov G., Milshina N.V., Mobarry C., Morris J., Mosrefi A.,  
Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclet J.M.,  
Palazololo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
Reiner B.C., Remington K., Saunders R.D., Scheeler F., Shen H.,  
Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
Svirskas R., Teetor C., Turner R., Venter E., Wang A.H., Wang X.,  
Wang Z.Y., Wassarman D.A., Weinstock G.M., Weissbach J.,  
Williams S.M., Woodager, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,  
Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
"The genome sequence of *Drosophila melanogaster*";  
*Science* 287:2185-2195(2000).  
[2]  
SEQUENCE FROM N.A.  
CP  
CElniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,  
Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,  
Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busan D.A.,  
Carlson J.W., Center A., Champs E., Davenport L.B., Dietz S.M.,  
Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,  
Ferriera S., Frise E., Galie R.F., Garg N.S., George R.A.,  
Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,  
Ibegwam C., Jalili M., Kruse D., Li P., Matteli B., Mosrefi A.,  
McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,  
Paclet J., Paragas V., Park S., Patel S., Pfeiffer B.,  
Phouanenavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,  
Stapleton M., Strong R., Svirskas R., Teetor C., Tyler D.,  
Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;  
"Sequencing of *Drosophila melanogaster* genome";  
Submitted (MAR-2000) to the ENBL/GenBank/DDBJ databases.  
[3]  
SEQUENCE FROM N.A.  
CP  
Mirra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,  
Hradecky P., Huang Y., Kaminker J.S., Prochak S.E., Smith C.D.,  
Tupy J.L., Bergman C., Bertram B., Carlson J.W., Celisner S.E.,  
Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,  
Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,  
Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield J.,  
Asburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;  
"Annotation of *Drosophila melanogaster* genome";  
Submitted (MAR-2000) to the ENBL/GenBank/DDBJ databases.  
[4]  
SEQUENCE FROM N.A.  
CP  
Adams M.D., Celisner S.E., Gibbs R.A., Rubin G.M., Venter C.J.;  
Submitted (MAR-2000) to the ENBL/GenBank/DDBJ databases.  
[5]  
SEQUENCE FROM N.A.  
CP  
FlyBase;  
DR  
EMBL Submitted (SEP-2002) to the ENBL/GenBank/DDBJ databases.  
DL  
EMBL: AE003753; AAFF56490.2; -  
DR  
FlyBase; FBgn0039354; lgr3.  
SQ  
SEQUENCE 765 AA; 87214 MW; 48FC480611f8BFD6 CRC64;  
Query Match 25.7%; Score 212.5; DB 5; Length 765;  
Best Local Similarity 35.9%; Pred. No. 2,2e-13;  
Matches 47; Conservative 24; Mismatches 55; Indels 5; Gaps 3;  
QY 3 IYSVATFLGINAAPTITIVSYSGMFVSQHQSATAETIRNQVKKEMLAKRFFFTVD 62  
DB 605 LGSAFYVLGVNLIIAMLIYALISWTWRTR-SATPL--TLIDCEFAVRFFFTVLD 660

```

QY 53 ALCWIPFYVAKPUSLQVEIPGTTISWVVGYSAINSLNPILYLTITRPPKEMH-RFW 121
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
D6 661 FLCWPPILVINKIIVFFNNYISDDIYAMLVFLVFLFSLASAVNPILYTTTPKYRNOIFLRGW 720
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 122 HNYQRKSMDS 132
      || | |
D6 721 KKITSRRKFAE 731
      || | |

```

```

RESULT 3
Q9DGC5 PRELIMINARY; PRT; 693 AA.
ID
AC Q9DGC5;
DT 01-WAR-2001 (TREMBlrel. 16, Created)
DT 01-WAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-WAR-2003 (TREMBlrel. 23, Last annotation update)
DE Gonadotropin receptor II.
GN TGT-RII.
OS Oreochromis niloticus (Nile tilapia) (Tilapia nilotica).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthopterygia; Acanthopterygii; Percomorpha; Perciformes; Labroidel;
OC Cichlidae; Oreochromis.
OX NCBI_TaxID=8128;
[1]
RP SEQUENCE FROM N.A.
RA Oba Y., Hirai T., Yoshiura Y., Yao Z., Nagahama Y.;
NT "Tilapia gonadotropin receptor II.";
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR EMBL; AB041763; BAB16107.1;
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR InterPro; IPR001611; LRR.
DR Pfam; PF00001; 7tm_1; 1.
DR Pfam; PF00560; LRR; 2.
DR PRINTS; PR00237; GPCR_RHODOPSN.
DR PROSITE; PS00237; G_PROTEIN_RECP_FL_1; 1.
DR PROSITE; PS0262; G_PROTEIN_RECP_FL_2; 1.
SQ G-protein coupled receptor; Receptor; Transmembrane.
KW SEQUENCE 693 AA; 777856 MW; C3BBEFD8ECFC8988 CRC64;

```

```

Query Match      18.6%; Score 154; DB 13; Length 693;
Best Local Similarity 34.5%; Pred. No. 2.2e-07;
Matches 40; Conservative 24; Mismatches 44; Indels 8; Gaps 4;

QY   1 AQIVSAIFIGINLAFFIIWFVSYGSMFYSHOSAITATEIRNOVKKEMILAKRFFIVF 60
    ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db   524 AQVTWAVLI-LNVAFIDWCYCICILSVHNPHESTR--RGDTK----IAERMAVLIF 576

QY   61 TDALCWLPFI-FVAKPLSLQVEIPGTITSNVVGYCAINSALNPILTYLTATRPKE 115
    ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db   577 TDFLCMAPISEFAAALRMELIWTWSHKILLIYPVINSCLNPFLLTITETRAFR 632

```

RESULT 4	
Q95YI6	
ID	Q95YI6
AC	PRELIMINARY; PRT; 1012 AA.
DC	Q95YI6;
DT	01-DEC-2001 (TrEMBLrel. 19, Created)
DT	01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT	01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE	Glycoprotein hormone receptor.
GN	AFGPHR.
OS	Asterina pectinifera (Starfish).
OC	Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Asterozoa;
OC	Asteroidea; Valvatacea; Valvatida; Asterinidae; Asterina.
OX	NCBI_TaxID=7594;
RN	[1]
RP	SEQUENCE FROM N.A.
RT	Mita M., Hirai T., Oba Y., Yoshikuni M., Nagahama Y.;
RA	"cDNA cloning and functional analysis of a novel member of the

RT glycoprotein hormone receptor family from a starfish Asterina  
 RL pectinifera.";  
 RT Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.  
 DR EMBL; AB061862; BAB68209.1; -  
 DR InterPro; IPR000276; GPCR\_Rhodpsn.  
 DR InterPro; IPR001611; LRR.  
 DR InterPro; IPR000372; LRR\_Nterm.  
 DR InterPro; IPR003591; LRR\_TYP.  
 DR Pfam; PF00001; 7tm\_1; 1.  
 DR Pfam; PF00560; LRR; 13.  
 DR PRINTS; PR00237; GPCRRHODPSN.  
 DR SMART; SM00013; LRRT; 1.  
 DR SMART; SM00369; LRR\_TYP; 5.  
 DR PROSITE; PS00237; G\_PROTEIN\_RECP\_FL\_1; 1.  
 DR PROSITE; PS0262; G\_PROTEIN\_RECP\_FL\_2; 1.  
 DR PROSITE; PS0506; LRR\_TYPICAL; 3.  
 KW G-protein coupled receptor; Receptor; Transmembrane.  
 SQ SEQUENCE 1012 AA; 112623 MW; 52A70E7A88C46E0A CRC64;

Query Match 17.9%; Score 148; DB 5; Length 1012;  
 Best Local Similarity 32.9%; Pred. No. 1.3e+06;  
 Matches 51; Conservative 27; Mismatches 51; Indels 26; Gaps 9

Gy 1 AQIYSVAIFLGILAAFIIVFSQSMFYSHQS-AITATEIRNOVKKEMILAKRFFTV 59  
 :::: |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:  
 Db 719 AKVY-VGSILLILAFVIINACYSIIYLAIQCSHAWNCDSR-----VAARRSLLV 769  
 Gy 60 FTDALCHWPI-----FVAKPGLSLQVEIPGTITSVVVICYSAINSAUNLPILYTTRPK 114  
 :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:  
 Db 770 FTDFAFWAPFAFTSLTAFLGLMLASDGAKVLATFVL----PLNSCANPELYTLTKQEK 825  
 Gy 115 ---EMIRWHNN--YRQRKSDSGIKRHMLHSS 144  
 :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:  
 Db 826 KCKTIYRSLSNRVFRQR-SM-SRSITLSGRHPS 858

RESULT 5  
 ID Q95YI7 PRELIMINARY; PRT: 1280 AA.  
 AC Q95YI7;  
 DT 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
 DE Glycoprotein hormone receptor.  
 GN ACPHR.  
 OS Asterina pectinifera (Starfish).  
 OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Asterozoa;  
 OC Asteroidea; Valvatacea; Valvutida; Asterinidae; Asterina.  
 OX NCBI\_TaxID=7594;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Mita M., Hirai T., Oba Y., Yoshikuni M., Nagahama Y.;  
 RT "cDNA cloning and functional analysis of a novel member of the  
 RT glycoprotein hormone receptor family from a starfish Asterina  
 RL pectinifera";  
 FL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.  
 DR EMBL; AB061861; BAB68208.1; -  
 DR InterPro; IPR000276; GPCR\_Rhodpsn.  
 DR InterPro; IPR001611; LRR.  
 DR InterPro; IPR000372; LRR\_Nterm.  
 DR InterPro; IPR003591; LRR\_TYP.  
 DR Pfam; PF00001; 7tm\_1; 1.  
 DR Pfam; PF00560; LRR; 13.  
 DR PRINTS; PR00237; GPCRRHODPSN.  
 DR SMART; SM00013; LRRT; 1.  
 DR SMART; SM00369; LRR\_TYP; 5.  
 DR PROSITE; PS00237; G\_PROTEIN\_RECP\_FL\_1; 1.  
 DR PROSITE; PS0262; G\_PROTEIN\_RECP\_FL\_2; 1.  
 DR PROSITE; PS0506; LRR\_TYPICAL; 3.



```

FT DISULFID 79 91 BY SIMILARITY.
FT DISULFID 86 104 BY SIMILARITY.
FT DISULFID 98 113 BY SIMILARITY.
FT DISULFID 118 131 BY SIMILARITY.
FT DISULFID 138 153 BY SIMILARITY.
FT DISULFID 158 170 BY SIMILARITY.
FT DISULFID 185 183 BY SIMILARITY.
FT DISULFID 177 194 BY SIMILARITY.
FT DISULFID 202 220 BY SIMILARITY.
FT DISULFID 214 230 BY SIMILARITY.
FT DISULFID 233 245 BY SIMILARITY.
FT DISULFID 240 258 BY SIMILARITY.
FT DISULFID 252 267 BY SIMILARITY.
FT DISULFID 274 291 BY SIMILARITY.
FT DISULFID 282 304 BY SIMILARITY.
FT DISULFID 298 316 BY SIMILARITY.
FT DISULFID 322 339 BY SIMILARITY.
FT DISULFID 334 352 BY SIMILARITY.
FT DISULFID 346 361 BY SIMILARITY.
FT DISULFID 367 379 BY SIMILARITY.
FT DISULFID 374 392 BY SIMILARITY.
FT DISULFID 386 401 BY SIMILARITY.
FT DISULFID 406 418 BY SIMILARITY.
FT DISULFID 413 431 BY SIMILARITY.
FT DISULFID 425 440 BY SIMILARITY.
FT DISULFID 416 458 BY SIMILARITY.
FT DISULFID 433 474 BY SIMILARITY.
FT DISULFID 465 483 BY SIMILARITY.
FT DISULFID 488 500 BY SIMILARITY.
FT DISULFID 495 513 BY SIMILARITY.
FT DISULFID 507 523 BY SIMILARITY.
FT CARBOHYD 87 87 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 166 166 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 269 269 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 318 318 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 482 482 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 502 502 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 571 571 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 618 618 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 624 624 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 685 685 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1115 AA; 145865 MW; E9BB01297ECE356C CRC64;

Query Match 25.1%; Score 207.5; DB 1; Length 1115;
Best Local Similarity 33.1%; Pred. No. 1.5e-11;
Matches 51; Conservative 25; Mismatches 45; Indels 33; Gaps 5;

QY 4 YSWAIFGLNLAFTIIIVFSGKMFYSV--HQSAITATEIRNOVKEMILAKRFFIIVF 60
DQ 938 YSGVFILLNLSFVLASSYLMWF-SVAKTRSAVRTAESKN----DNAMARRMTLIW 992
QY 61 TDALCWIPFVAKPLSLQVEIPGTTISWVIGYSAINSLNPLIYLTLPTRP----- 113
DQ 993 TDFCCWVPIVLGFGVSLAGARADDQYANIVFVPLASATNPVIYTLSTAPFLGNVSR 1052
QY 114 -----KEMTHRP-----WHNTRQKS 129
DQ 1053 ANRFKSFTHSTGTGTHKSHYVDDGTGTHSYCEKKS 1086

RESULT 5
FSHR_SHEEP
ID FSHR_SHEEP STANDARD; PRT; 695 AA.
AC P35379; Q28573; Q28574; Q9TS19;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Follicle stimulating hormone receptor precursor (FSH-R) (Follitropin
  receptor).
GN FSHR.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

```

```

OC Boridae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RN RP SEQUENCE FROM N.A. (ISOFORM FSH-R1).
RC TISSUE=Testis;
RX MEDLINE=93351750; PubMed=8394255;
RA Yarney T.A., Sairam M.R., Khan H., Ravindranath N., Payne S.,
RA Seidman N.G.;
RT "Molecular cloning and expression of the ovine testicular follicle
  stimulating hormone receptor.";
RL Mol. Cell. Endocrinol. 93:219-226(1993).
RN [2]
RN RP SEQUENCE FROM N.A. (ISOFORMS FSH-R4 AND FSH-R3).
RC STRAIN=Dorset-Leicester-Suffolk 1; TISSUE=Testis;
RX MEDLINE=93176195; PubMed=8493338;
RA Khan H., Yarney T.A., Sairam M.R.;
RT "Cloning of alternately spliced mRNA transcripts coding for variants
  of ovine testicular follitropin receptor lacking the G protein
  coupling domains.";
RL Biochem. Biophys. Res. Commun. 190:888-894(1993).
RN [3]
RN RP SEQUENCE FROM N.A. (ISOFORM FSH-R2), AND CHARACTERIZATION.
RC STRAIN=Dorset-Leicester-Suffolk 1; TISSUE=Testis;
RX MEDLINE=98031015; PubMed=9364440;
RA Yarney T.A., Jiang L., Khan H., MacDonald E.A., Laird D.W.,
RA Sairam M.R.;
RT "Molecular cloning, structure, and expression of a testicular
  follitropin receptor with selective alteration in the carboxy terminus
  that affects signaling function.";
RL Mol. Reprod. Dev. 48:458-470(1997).
RN [4]
RN RP SEQUENCE FROM N.A. (ISOFORM FSH-R3), AND CHARACTERIZATION.
RC STRAIN=Dorset-Leicester-Suffolk 1; TISSUE-Ovary;
RX MEDLINE=20391225; PubMed=10527866;
RA Babu P.S., Jiang L., Sairam A.M., Toney R.M., Sairam M.R.;
RT "Structural features and expression of an alternatively spliced growth
  factor type I receptor for follitropin signaling in the developing
  ovary.";
RL Mol. Cell Biol. Res. Commun. 2:21-27(1999).
RN [5]
RN RP SEQUENCE OF 1-51 FROM N.A.
RX MEDLINE=98031017; PubMed=9364442;
RA Sairam M.R., Subbarayan V.S.R.;
RT "Characterization of the 5' flanking region and potential control
  elements of the ovine follitropin receptor gene.";
RL Mol. Reprod. Dev. 48:480-487(1997).
CC -1- FUNCTION: Receptor for follicle stimulating hormone. The activity
  of isoform FSH-R1 is mediated by G proteins which activate
  adenylyate cyclase. Isoforms FSH-R2 and FSH-R3 also bind FSH, but
  this does not result in activation of adenylyate cyclase. Isoform
  FSH-R3 may be involved in calcium signaling.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Plasma membrane
  (isoforms FSH-R1 and FSH-R2); Cell surface (isoform FSH-R3).
CC -1- ALTERNATIVE PRODUCTS:
  Event=Alternative splicing; Named isoforms=4;
  Name=FSH-R1;
  IsoId=P35379-1; Sequence=Displayed;
  Name=FSH-R2;
  IsoId=P35379-2; Sequence=VSP_001959, VSP_001960;
  Name=FSH-R3;
  IsoId=P35379-3; Sequence=VSP_001957, VSP_001958;
  Name=FSH-R4;
  IsoId=P35379-4; Sequence=VSP_001955, VSP_001956;
CC -1- TISSUE SPECIFICITY: Isoform FSH-R3 is expressed in ovary and
  testis, but not in kidney.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
  FSH/LSH/TSH SUBFAMILY.
CC -1- SIMILARITY: Contains 6 leucine-rich (LRR) repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way

```

CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 CC EMBL; L07302; AAA31525.1; -;  
 CC EMBL; L12766; AAA31523.1; -;  
 CC EMBL; L12767; AAA31524.1; -;  
 CC EMBL; L36115; AAK70667.1; -;  
 CC EMBL; AJ131735; CAA10495.1; -;  
 CC EMBL; AF090438; AAC61749.1; -;  
 CC FIR; JCI493; JCI493.  
 CC HSP; P23945; IXUN.  
 CC InterPro; IPR000276; GPCR\_Rhodopsn.  
 CC InterPro; IPR001611; LRR.  
 CC InterPro; IPR000372; LRR\_Nterm.  
 CC Pfam; PF00001; 7tm\_1; 1.  
 CC Pfam; PF00560; LRR; 5.  
 CC Pfam; PF01462; LRRNT; 1.  
 CC PRINTS; PR00237; GPCRHOPOSPN.  
 CC SMART; SM00013; LRRNT; 1.  
 CC PROSITE; PS00237; G-PROTEIN\_RECF1\_1; 1.  
 CC PROSITE; PS0262; G-PROTEIN\_RECF2\_1; 1.  
 CC G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;  
 KW Phosphorylation; Repeat; Leucine-rich repeat; Alternative splicing.  
 FT SIGNAL  
 FT CHAIN 1  
 FT DOMAIN 18 695  
 FT TRANSMEM 18 366  
 FT TRANSMEM 367 387  
 FT TRANSMEM 388 398  
 FT TRANSMEM 399 421  
 FT TRANSMEM 422 443  
 FT TRANSMEM 444 465  
 FT TRANSMEM 466 485  
 FT TRANSMEM 486 508  
 FT TRANSMEM 509 528  
 FT TRANSMEM 529 550  
 FT TRANSMEM 551 573  
 FT TRANSMEM 574 597  
 FT TRANSMEM 598 608  
 FT TRANSMEM 609 630  
 FT TRANSMEM 631 695  
 FT REPEAT 44 68  
 FT REPEAT 69 93  
 FT REPEAT 119 143  
 FT REPEAT 170 192  
 FT REPEAT 193 216  
 FT REPEAT 218 240  
 FT REPEAT 240 259  
 FT DISULFID 442 517  
 FT CARBOHYD 191 191  
 FT CARBOHYD 199 199  
 FT CARBOHYD 293 293  
 FT VARSPPLIC 126 133  
 FT VARSPPLIC 135 695  
 FT VARSPPLIC 224 259  
 FT VARSPPLIC 260 695  
 FT VARSPPLIC 643 670  
 FT VARSPPLIC 671 695  
 FT VARSPPLIC 695 AA; 78237 MW; FBF75D89D8C0D4B CRC64;  
 FT SEQUENCE 17.4%; Score 143.5; DB 1; Length 695;  
 Query Match  
 Best Local Similarity 27.2%; Pred. No. 7.7e-06;  
 Matches 44; Conservative 33; Mismatches 59; Indels 27; Gaps 6;  
 QY 1 AQIYSVALFLGINLAFFIIIVSYGSMFYSHQSAITATEIRNOVKEMILAKRFFIIVF 60

Db 527 SOLYMSL-LVLNLVAFWICGCTHYITVFNPTSS-----SSDTKIARWALIF 579  
 QY 61 TDALCWIPF-FVAKPLSLQVPEIGTITSVVGVYSAINALNPILYITLTTPPKF----- 115  
 Db 580 TDFCLMAPISFAISAKVPLITVSKSLILVLFYFIPINSCANPFYIAITFRFRDFFI 639  
 QY 116 MIHFW-----HNYRQKSMDSKGIKHEMLHSHSGKCGHC 151  
 Db 640 LLSKFCYEVQAOYRSETSFTH-----HNFHPNGHC 672  
 -----  
 RESULT 6  
 GLHR ANTEL  
 ID GLHR ANTEL STANDARD; PRT; 925 AA.  
 AC P35409; 1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Probable glycoprotein hormone G-protein coupled receptor precursor.  
 OS Anthopleura elegantissima (Sea anemone).  
 CC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Actiniaria;  
 CC Nynanthaeae; Actiniidae; Anthopleura.  
 CC NCBI\_TaxID=6110;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORMS LONG AND SHORT).  
 RX MEDLINE=94107299; PubMed=8280121;  
 RA Nothacker H.-P., Grimmelikhuijzen C.J.P.;  
 RT "Molecular cloning of a novel, putative G protein-coupled receptor  
 from sea anemones structurally related to members of the FSH, TSH,  
 LH/CG receptor family from mammals";  
 RT Biochem. Biophys. Res. Commun. 197:1062-1069(1993).  
 CC -!- FUNCTION: PROBABLE RECEPTOR FOR A GLYCOPROTEIN HORMONE.  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=2;  
 CC Name=Long;  
 CC IsoId=P35409-1; Sequence=Displayed;  
 CC Name=Short;  
 CC IsoId=P35409-2; Sequence=VSP\_001983;  
 CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.  
 CC FSH/LSH/TSH SUBFAMILY.  
 -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 CC EMBL; 228332; CAA82186.1; -;  
 CC FIR; JC2033; JC2033.  
 CC InterPro; IPR000276; GPCR\_Rhodopsn.  
 CC InterPro; IPR001611; LRR.  
 CC InterPro; IPR003591; LRR\_typ.  
 CC Pfam; PF00001; 7tm\_1; 1.  
 CC Pfam; PF00560; LRR; 6.  
 CC PRINTS; PR00237; GPCRHOPOSPN.  
 CC PROSITE; PS00237; G-PROTEIN\_RECF1\_1; 1.  
 CC PROSITE; PS0262; G-PROTEIN\_RECF2\_1; 1.  
 CC G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;  
 KW Alternative splicing; Repeat.  
 FT SIGNAL 1 27  
 FT CHAIN 28 925  
 FT PROBABLE GLYCOPROTEIN HORMONE G-PROTEIN  
 FT COUPLED RECEPTOR.  
 FT EXTRACELLULAR (POTENTIAL).  
 FT DOMAIN 28 529  
 FT TRANSMEM 530 551  
 FT DOMAIN 552 561  
 FT TRANSMEM 562 584  
 FT DOMAIN 585 606  
 FT TRANSMEM 607 628  
 FT DOMAIN 629 651  
 FT CYTOPLASMIC (POTENTIAL).

```

FT TRANSMEM 652 673 4 (POTENTIAL).
FT DOMAIN 674 691 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 692 712 5 (POTENTIAL).
FT DOMAIN 713 739 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 740 763 6 (POTENTIAL).
FT DOMAIN 764 774 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 775 795 7 (POTENTIAL).
FT DOMAIN 796 925 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 333 461 5 X APPROXIMATE TANDEM REPEATS.
FT REPEAT 333 349 1 (INCOMPLETE).
FT REPEAT 350 384 2.
FT REPEAT 385 419 3.
FT REPEAT 420 453 4.
FT REPEAT 454 461 5 (INCOMPLETE).
FT CARBOHYD 61 61 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 152 152 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 212 212 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 435 435 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 495 495 N-LINKED (GLCNAC. .) (POTENTIAL).
FT VARSPLIC 235 925 Missing (in isoform Short).
SQ SEQUENCE 925 AA; 100059 MW; D03A256368452FED CRC64;

Query Match 17.3%; Score 143; DB 1; Length 925;
Best Local Similarity 33.3%; Pred. No. 1.2e-05;
Matches 37; Conservative 21; Mismatches 47; Indels 6; Gaps 3;

QY 6 VALFGLNLAAPLIIVFSGMFSYVHQSALTATEIRNOVKK-EMILAKREFEIVFDAL 64
DB 694 VALLVNGASFIWYLSMLV- ----VSGDMGAPKRNDSKVAKRAILVFTDKL 749
QY 65 CWIPI-FVAKPLSLQVEIPGTTISWVIGYSAINSLNPLYLTPTRPK 114
DB 750 CWAPIAFFGLLAAGQTLIVTQSKILLVFEFPINSICNPFLYAFETKPK 800

RESULT 7
FSHR_MACFA STANDARD; PRT; 695 AA.
AC P32212;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Follicle stimulating hormone receptor precursor (FSH-R) (Folliotropin receptor).
DE FSHR.
GN FSHR.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=94071854; PubMed=7504463;
RA Gromoll J., Dankbar B., Sharma R.S., Nieschlag E.;
RT "Molecular cloning of the testicular follicle stimulating hormone receptor of the non human primate Macaca fascicularis and identification of multiple transcripts in the testis.";
RL Biochem. Biophys. Res. Commun. 196;1066-1072(1993).
CC -!- FUNCTION: RECEPTOR FOR FOLLICLE STIMULATING HORMONE. THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE ADENYLATE CYCLASE.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC FSH/LSH/TSH SUBFAMILY.
CC -!- SIMILARITY: Contains 6 leucine-rich (LRR) repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial

```

```

CC entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X74454; CA52463.1; -.
DR PIR; J08998; J08998.
DR HSSP; P23945; LXUN.
DR InterPro; IPR00276; GPCR_Rhodpsn.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR00372; LRR_Nterm.
DR Pfam; PF00001; 7cm_1; 1.
DR Pfam; PF00360; LRR; 3.
DR PRINTS; PR0037; GPCRHDOPSN.
DR SMART; SM0013; LRRNT; 1.
DR PROSITE; PS00337; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS0362; G_PROTEIN_RECEP_F1_2; 1.
KW G-protein coupled receptor; transmembrane; glycoprotein; signal;
KW phosphorylation; repeat; leucine-rich repeat.
FT SIGNAL 1 17 POTENTIAL.
FT CHAIN 18 695 FOLLICLE STIMULATING HORMONE RECEPTOR.
FT DOMAIN 18 366 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 367 387 1 (POTENTIAL).
FT DOMAIN 388 398 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 399 421 2 (POTENTIAL).
FT DOMAIN 422 443 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 444 465 3 (POTENTIAL).
FT DOMAIN 466 485 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 486 508 4 (POTENTIAL).
FT DOMAIN 509 528 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 529 550 5 (POTENTIAL).
FT DOMAIN 551 573 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 574 597 6 (POTENTIAL).
FT DOMAIN 598 608 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 609 630 7 (POTENTIAL).
FT DOMAIN 631 695 CYTOPLASMIC (POTENTIAL).
FT REPEAT 44 68 LRR 1.
FT REPEAT 69 93 LRR 2.
FT REPEAT 119 143 LRR 3.
FT REPEAT 170 192 LRR 4.
FT REPEAT 193 216 LRR 5.
FT REPEAT 218 240 LRR 6.
FT DISULFID 442 517 BY SIMILARITY.
FT CARBOHYD 191 191 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 199 199 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 293 293 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 318 318 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 695 AA; 78343 MW; 0D60A233729B5250 CRC64;

Query Match 17.1%; Score 141.5; DB 1; Length 695;
Best Local Similarity 26.5%; Pred. No. 1.2e-05;
Matches 44; Conservative 32; Mismatches 63; Indels 27; Gaps 6;

QY 1 AQIYVAIFGLNLAAPLIIVFSGMFSYVHQSALTATEIRNOVKKEMILAKREFEIVF 60
DB 527 SOLIYMSL-LVLNVLAFVVICGCTHYLIVRNPIVSS-----SSDRIAKRMALIF 579
QY 61 TDALCWIPIL-FVAKPLSLQVEIPGTTISWVIGYSAINSLNPLYLTPTRPK----- 115
DB 580 TDFLCVAPISFPAISASLKVPLITVSKAKILLVLFVPIVNSCANPFLYAIETKNERDFT 639
QY 116 MHRFW-----HNYRQKSDSKGIRKMKLHSHSSGGKCGHCRCH 155
DB 640 LLSKFCGYEMOQIYRTSTSTA-----HNSHPNGHCHSSAH 676

RESULT 8
LSHR_PIG STANDARD; PRT; 696 AA.
ID LSHR_PIG
AC P16582;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Lutropin-choriogonadotropic hormone receptor precursor (LH/CG-R)

```





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 25, 2003, 14:54:41 ; Search time 11 Seconds  
(without alignments)  
671.199 Million cell updates/sec

Title: US-10-049-568-2

Perfect score: 826

Sequence: 1 AQYVAIFLGINLAFTII.....HMLHSSGGKCHCRRCHLS 157

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	638	77.2	757	1	LGR7_HUMAN
2	409	49.5	737	1	LGR8_MOUSE
3	392.5	47.5	754	1	LGR8_HUMAN
4	207.5	25.1	1115	1	GPCR_LYMST
5	143.5	17.4	695	1	FSHR_SHEEP
6	143	17.3	923	1	GLHR_ANGEL
7	141.5	17.1	695	1	FSHR_MACFA
8	140.5	17.0	696	1	LSHR_PIG
9	139.5	16.9	692	1	FSHR_RAT
10	139.5	16.9	695	1	FSHR_BOVIN
11	139	16.8	676	1	LSHR_CALJA
12	136.5	16.5	695	1	FSHR_PIG
13	134	16.2	700	1	LSHR_MOUSE
14	134	16.2	700	1	LSHR_RAT
15	132.5	16.0	692	1	FSHR_MOUSE
16	132	16.0	701	1	LSHR_BOVIN
17	131.5	15.9	694	1	FSHR_HORSE
18	131	15.9	699	1	LSHR_HUMAN
19	131	15.9	699	1	FSHR_EQUAS
20	129	15.6	687	1	FSHR_MOUSE
21	128.5	15.6	319	1	AA3R_MOUSE
22	128	15.5	320	1	AA3R_RAT
23	127	15.4	764	1	TSHR_SHEEP
24	125.5	15.2	695	1	FSHR_HUMAN
25	124	15.0	338	1	P2YX_HUMAN
26	124	15.0	763	1	TSHR_BOVIN
27	124	15.0	764	1	TSHR_RAT
28	123	14.9	340	1	AA2B_CHICK
29	123	14.9	366	1	LSHR_CHICK
30	122	14.8	764	1	TSHR_CANFA
31	121	14.6	764	1	TSHR_MOUSE
32	119.5	14.5	332	1	AA2B_RAT
33	119.5	14.5	378	1	EDG3_HUMAN

RESULT 1  
LGR7\_HUMAN  
ID LGR7\_HUMAN STANDARD; PRT; 757 AA.  
AC Q9HEX9;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Relaxin receptor 1 (Leucine-rich repeat-containing G protein-coupled receptor 7).  
DE receptor 7).  
GN LGR7.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OC NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), AND MUTAGENESIS OF ASP-637.  
RX MEDLINE-20338592; PubMed-10935549;  
RA Hsu S.Y., Kudo M., Chen T., Nakabayashi K., Bhalla A., van der Spek P.J., van Duin M., Hsueh A.J.W.;  
RT "The three subfamilies of leucine-rich repeat-containing G protein-coupled receptors (LGR): identification of LGR6 and LGR7 and the signalling mechanism for LGR7";  
RT Mol. Endocrinol. 14:1257-1271(2000).  
RL [2]  
RN CHARACTERIZATION.  
RP MEDLINE-21408393; PubMed-11517286;  
RX Bartsch O., Bartlick B., Iwell R.;  
RA "Relaxin signalling links tyrosine phosphorylation to phosphodiesterase and adenylyl cyclase activity";  
RT Mol. Hum. Reprod. 7:799-809(2001).  
RL  
CC -!- FUNCTION: Receptor for relaxin. The activity of this receptor is mediated by G proteins leading to stimulation of adenylyl cyclase and an increase of cAMP. Binding of the ligand may also activate a tyrosine kinase pathway that inhibits the activity of a phosphodiesterase that degrades cAMP.  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
CC -!- ALTERNATIVE PRODUCTS:  
CC Event-Alternative splicing: Named isoforms=2;  
CC Name=1;  
CC Name=2;  
CC IsoId=Q9HEX9-1; Sequence-Displayed;  
CC IsoId=Q9HEX9-2; Sequence-VSP\_001384;  
CC -!- TISSUE SPECIFICITY: Expressed in the brain, kidney, testis, placenta, uterus, ovary, adrenal, prostate, skin and heart. Not detected in spleen.  
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.  
CC -!- SIMILARITY: Contains 1 LDL-receptor class A domain.  
CC -!- SIMILARITY: Contains 10 leucine-rich (LRR) repeats.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

34 119.5 14.5 417 1 A1AB\_CANFA  
35 119 14.4 519 1 A1AB\_HUMAN  
36 118.5 14.3 317 1 AA3R\_SHEEP  
37 118.5 14.3 338 1 P2YX\_MOUSE  
38 118 14.3 332 1 AA2B\_MOUSE  
39 117.5 14.2 514 1 A1AB\_MOUSE  
40 117.5 14.2 515 1 A1AB\_MESAU  
41 117.5 14.2 515 1 A1AB\_RAT  
42 117.5 14.2 522 1 NFF2\_HUMAN  
43 117 14.2 326 1 AALR\_BOVIN  
44 117 14.2 326 1 AALR\_CAVPO  
45 117 14.2 332 1 AA2B\_HUMAN

## ALIGNMENTS

P11615 canis faml  
P35368 homo sapien  
P35342 ovis aries  
Q9esg6 mus musculu  
Q60614 mus musculu  
P97717 mus musculu  
P18841 mesocricetu  
P15823 rattus norv  
Q9yxs3 homo sapien  
P28190 bos taurus  
P47745 cavia porce  
P29275 homo sapien

```

CC DR EMBL: AE190500; AAG17167.1; -.
DR MM; 608634; -.
DR HSP; P01130; ILDR.
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR DR InterPro: IPR002172; LDL_receptor_A.
DR DR InterPro: IPR001611; LRR.
DR DR InterPro: IPR000372; LRR_Nterm.
DR DR InterPro: IPR003591; LRR_typ.
DR DR Pfam; PF00001; 7tm_1; 1.
DR DR Pfam; PF00057; LDL_recept_a; 1.
DR DR Pfam; PF00560; LRR; 8.
DR DR PRINTS; PR00237; GPCRHOOPS.
DR DR SMART; SM00192; LDLa; 1.
DR DR SMART; SM00369; LRR_TVP; 2.
DR DR SMART; SM00013; LRRNT; 1.
DR DR PROSITE; PS00237; G-PROTEIN_RECP_F1_1; FALSE_NEG.
DR DR PROSITE; PS0262; G-PROTEIN_RECP_F1_2; 1.
DR DR PROSITE; PS01209; LDLRA_1; 1.
DR DR PROSITE; PS00068; LDLRA_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Repeat;
KW Leucine-rich repeat; Alternative splicing.
FT DOMAIN 1 409
FT TRANSSEM 410 430
FT DOMAIN 431 443
FT TRANSSEM 444 464
FT DOMAIN 465 486
FT TRANSSEM 487 507
FT DOMAIN 508 527
FT TRANSSEM 528 548
FT DOMAIN 549 577
FT TRANSSEM 578 598
FT DOMAIN 599 629
FT TRANSSEM 630 650
FT DOMAIN 651 671
FT TRANSSEM 672 692
FT DOMAIN 693 713
FT TRANSSEM 714 734
FT DOMAIN 735 755
FT TRANSSEM 756 776
FT DOMAIN 777 797
FT TRANSSEM 798 818
FT DOMAIN 819 839
FT TRANSSEM 840 860
FT DOMAIN 861 881
FT TRANSSEM 882 902
FT DOMAIN 903 923
FT TRANSSEM 924 944
FT DOMAIN 945 965
FT TRANSSEM 966 986
FT DOMAIN 987 1007
FT TRANSSEM 1008 1028
FT DOMAIN 1029 1049
FT TRANSSEM 1050 1070
FT DOMAIN 1071 1091
FT TRANSSEM 1092 1112
FT DOMAIN 1113 1133
FT TRANSSEM 1134 1154
FT DOMAIN 1155 1175
FT TRANSSEM 1176 1196
FT DOMAIN 1197 1217
FT TRANSSEM 1218 1238
FT DOMAIN 1239 1259
FT TRANSSEM 1260 1280
FT DOMAIN 1281 1301
FT TRANSSEM 1302 1322
FT DOMAIN 1323 1343
FT TRANSSEM 1344 1364
FT DOMAIN 1365 1385
FT TRANSSEM 1386 1406
FT DOMAIN 1407 1427
FT TRANSSEM 1428 1448
FT DOMAIN 1449 1469
FT TRANSSEM 1470 1490
FT DOMAIN 1491 1511
FT TRANSSEM 1512 1532
FT DOMAIN 1533 1553
FT TRANSSEM 1554 1574
FT DOMAIN 1575 1595
FT TRANSSEM 1596 1616
FT DOMAIN 1617 1637
FT TRANSSEM 1638 1658
FT DOMAIN 1659 1679
FT TRANSSEM 1680 1700
FT DOMAIN 1701 1721
FT TRANSSEM 1722 1742
FT DOMAIN 1743 1763
FT TRANSSEM 1764 1784
FT DOMAIN 1785 1805
FT TRANSSEM 1806 1826
FT DOMAIN 1827 1847
FT TRANSSEM 1848 1868
FT DOMAIN 1869 1889
FT TRANSSEM 1890 1910
FT DOMAIN 1911 1931
FT TRANSSEM 1932 1952
FT DOMAIN 1953 1973
FT TRANSSEM 1974 1994
FT DOMAIN 1995 2015
FT TRANSSEM 2016 2036
FT DOMAIN 2037 2057
FT TRANSSEM 2058 2078
FT DOMAIN 2079 2099
FT TRANSSEM 2100 2120
FT DOMAIN 2121 2141
FT TRANSSEM 2142 2162
FT DOMAIN 2163 2183
FT TRANSSEM 2184 2204
FT DOMAIN 2205 2225
FT TRANSSEM 2226 2246
FT DOMAIN 2247 2267
FT TRANSSEM 2268 2288
FT DOMAIN 2289 2309
FT TRANSSEM 2310 2330
FT DOMAIN 2331 2351
FT TRANSSEM 2352 2372
FT DOMAIN 2373 2393
FT TRANSSEM 2394 2414
FT DOMAIN 2415 2435
FT TRANSSEM 2436 2456
FT DOMAIN 2457 2477
FT TRANSSEM 2478 2498
FT DOMAIN 2499 2519
FT TRANSSEM 2520 2540
FT DOMAIN 2541 2561
FT TRANSSEM 2562 2582
FT DOMAIN 2583 2603
FT TRANSSEM 2604 2624
FT DOMAIN 2625 2645
FT TRANSSEM 2646 2666
FT DOMAIN 2667 2687
FT TRANSSEM 2688 2708
FT DOMAIN 2709 2729
FT TRANSSEM 2730 2750
FT DOMAIN 2751 2771
FT TRANSSEM 2772 2792
FT DOMAIN 2793 2813
FT TRANSSEM 2814 2834
FT DOMAIN 2835 2855
FT TRANSSEM 2856 2876
FT DOMAIN 2877 2897
FT TRANSSEM 2898 2918
FT DOMAIN 2919 2939
FT TRANSSEM 2940 2960
FT DOMAIN 2961 2981
FT TRANSSEM 2982 3002
FT DOMAIN 3003 3023
FT TRANSSEM 3024 3044
FT DOMAIN 3045 3065
FT TRANSSEM 3066 3086
FT DOMAIN 3087 3107
FT TRANSSEM 3108 3128
FT DOMAIN 3129 3149
FT TRANSSEM 3150 3170
FT DOMAIN 3171 3191
FT TRANSSEM 3192 3212
FT DOMAIN 3213 3233
FT TRANSSEM 3234 3254
FT DOMAIN 3255 3275
FT TRANSSEM 3276 3296
FT DOMAIN 3297 3317
FT TRANSSEM 3318 3338
FT DOMAIN 3339 3359
FT TRANSSEM 3360 3380
FT DOMAIN 3381 3401
FT TRANSSEM 3402 3422
FT DOMAIN 3423 3443
FT TRANSSEM 3444 3464
FT DOMAIN 3465 3485
FT TRANSSEM 3486 3506
FT DOMAIN 3507 3527
FT TRANSSEM 3528 3548
FT DOMAIN 3549 3569
FT TRANSSEM 3570 3590
FT DOMAIN 3591 3611
FT TRANSSEM 3612 3632
FT DOMAIN 3633 3653
FT TRANSSEM 3654 3674
FT DOMAIN 3675 3695
FT TRANSSEM 3696 3716
FT DOMAIN 3717 3737
FT TRANSSEM 3738 3758
FT DOMAIN 3759 3779
FT TRANSSEM 3780 3800
FT DOMAIN 3801 3821
FT TRANSSEM 3822 3842
FT DOMAIN 3843 3863
FT TRANSSEM 3864 3884
FT DOMAIN 3885 3905
FT TRANSSEM 3906 3926
FT DOMAIN 3927 3947
FT TRANSSEM 3948 3968
FT DOMAIN 3969 3989
FT TRANSSEM 3990 4010
FT DOMAIN 4011 4031
FT TRANSSEM 4032 4052
FT DOMAIN 4053 4073
FT TRANSSEM 4074 4094
FT DOMAIN 4095 4115
FT TRANSSEM 4116 4136
FT DOMAIN 4137 4157
FT TRANSSEM 4158 4178
FT DOMAIN 4179 4199
FT TRANSSEM 4200 4220
FT DOMAIN 4221 4241
FT TRANSSEM 4242 4262
FT DOMAIN 4263 4283
FT TRANSSEM 4284 4304
FT DOMAIN 4305 4325
FT TRANSSEM 4326 4346
FT DOMAIN 4347 4367
FT TRANSSEM 4368 4388
FT DOMAIN 4389 4409
FT TRANSSEM 4410 4430
FT DOMAIN 4431 4451
FT TRANSSEM 4452 4472
FT DOMAIN 4473 4493
FT TRANSSEM 4494 4514
FT DOMAIN 4515 4535
FT TRANSSEM 4536 4556
FT DOMAIN 4557 4577
FT TRANSSEM 4578 4598
FT DOMAIN 4599 4619
FT TRANSSEM 4620 4640
FT DOMAIN 4641 4661
FT TRANSSEM 4662 4682
FT DOMAIN 4683 4703
FT TRANSSEM 4704 4724
FT DOMAIN 4725 4745
FT TRANSSEM 4746 4766
FT DOMAIN 4767 4787
FT TRANSSEM 4788 4808
FT DOMAIN 4809 4829
FT TRANSSEM 4830 4850
FT DOMAIN 4851 4871
FT TRANSSEM 4872 4892
FT DOMAIN 4893 4913
FT TRANSSEM 4914 4934
FT DOMAIN 4935 4955
FT TRANSSEM 4956 4976
FT DOMAIN 4977 4997
FT TRANSSEM 4998 5018
FT DOMAIN 5019 5039
FT TRANSSEM 5040 5060
FT DOMAIN 5061 5081
FT TRANSSEM 5082 5102
FT DOMAIN 5103 5123
FT TRANSSEM 5124 5144
FT DOMAIN 5145 5165
FT TRANSSEM 5166 5186
FT DOMAIN 5187 5207
FT TRANSSEM 5208 5228
FT DOMAIN 5229 5249
FT TRANSSEM 5250 5270
FT DOMAIN 5271 5291
FT TRANSSEM 5292 5312
FT DOMAIN 5313 5333
FT TRANSSEM 5334 5354
FT DOMAIN 5355 5375
FT TRANSSEM 5376 5396
FT DOMAIN 5397 5417
FT TRANSSEM 5418 5438
FT DOMAIN 5439 5459
FT TRANSSEM 5460 5480
FT DOMAIN 5481 5501
FT TRANSSEM 5502 5522
FT DOMAIN 5523 5543
FT TRANSSEM 5544 5564
FT DOMAIN 5565 5585
FT TRANSSEM 5586 5606
FT DOMAIN 5607 5627
FT TRANSSEM 5628 5648
FT DOMAIN 5649 5669
FT TRANSSEM 5670 5690
FT DOMAIN 5691 5711
FT TRANSSEM 5712 5732
FT DOMAIN 5733 5753
FT TRANSSEM 5754 5774
FT DOMAIN 5775 5795
FT TRANSSEM 5796 5816
FT DOMAIN 5817 5837
FT TRANSSEM 5838 5858
FT DOMAIN 5859 5879
FT TRANSSEM 5880 5900
FT DOMAIN 5901 5921
FT TRANSSEM 5922 5942
FT DOMAIN 5943 5963
FT TRANSSEM 5964 5984
FT DOMAIN 5985 6005
FT TRANSSEM 6006 6026
FT DOMAIN 6027 6047
FT TRANSSEM 6048 6068
FT DOMAIN 6069 6089
FT TRANSSEM 6090 6110
FT DOMAIN 6111 6131
FT TRANSSEM 6132 6152
FT DOMAIN 6153 6173
FT TRANSSEM 6174 6194
FT DOMAIN 6195 6215
FT TRANSSEM 6216 6236
FT DOMAIN 6237 6257
FT TRANSSEM 6258 6278
FT DOMAIN 6279 6299
FT TRANSSEM 6300 6320
FT DOMAIN 6321 6341
FT TRANSSEM 6342 6362
FT DOMAIN 6363 6383
FT TRANSSEM 6384 6404
FT DOMAIN 6405 6425
FT TRANSSEM 6426 6446
FT DOMAIN 6447 6467
FT TRANSSEM 6468 6488
FT DOMAIN 6489 6509
FT TRANSSEM 6510 6530
FT DOMAIN 6531 6551
FT TRANSSEM 6552 6572
FT DOMAIN 6573 6593
FT TRANSSEM 6594 6614
FT DOMAIN 6615 6635
FT TRANSSEM 6636 6656
FT DOMAIN 6657 6677
FT TRANSSEM 6678 6698
FT DOMAIN 6699 6719
FT TRANSSEM 6720 6740
FT DOMAIN 6741 6761
FT TRANSSEM 6762 6782
FT DOMAIN 6783 6803
FT TRANSSEM 6804 6824
FT DOMAIN 6825 6845
FT TRANSSEM 6846 6866
FT DOMAIN 6867 6887
FT TRANSSEM 6888 6908
FT DOMAIN 6909 6929
FT TRANSSEM 6930 6950
FT DOMAIN 6951 6971
FT TRANSSEM 6972 6992
FT DOMAIN 6993 7013
FT TRANSSEM 7014 7034
FT DOMAIN 7035 7055
FT TRANSSEM 7056 7076
FT DOMAIN 7077 7097
FT TRANSSEM 7098 7118
FT DOMAIN 7119 7139
FT TRANSSEM 7140 7160
FT DOMAIN 7161 7181
FT TRANSSEM 7182 7202
FT DOMAIN 7203 7223
FT TRANSSEM 7224 7244
FT DOMAIN 7245 7265
FT TRANSSEM 7266 7286
FT DOMAIN 7287 7307
FT TRANSSEM 7308 7328
FT DOMAIN 7329 7349
FT TRANSSEM 7350 7370
FT DOMAIN 7371 7391
FT TRANSSEM 7392 7412
FT DOMAIN 7413 7433
FT TRANSSEM 7434 7454
FT DOMAIN 7455 7475
FT TRANSSEM 7476 7496
FT DOMAIN 7497 7517
FT TRANSSEM 7518 7538
FT DOMAIN 7539 7559
FT TRANSSEM 7560 7580
FT DOMAIN 7581 7601
FT TRANSSEM 7602 7622
FT DOMAIN 7623 7643
FT TRANSSEM 7644 7664
FT DOMAIN 7665 7685
FT TRANSSEM 7686 7706
FT DOMAIN 7707 7727
FT TRANSSEM 7728 7748
FT DOMAIN 7749 7769
FT TRANSSEM 7770 7790
FT DOMAIN 7791 7811
FT TRANSSEM 7812 7832
FT DOMAIN 7833 7853
FT TRANSSEM 7854 7874
FT DOMAIN 7875 7895
FT TRANSSEM 7896 7916
FT DOMAIN 7917 7937
FT TRANSSEM 7938 7958
FT DOMAIN 7959 7979
FT TRANSSEM 7980 8000
FT DOMAIN 8001 8021
FT TRANSSEM 8022 8042
FT DOMAIN 8043 8063
FT TRANSSEM 8064 8084
FT DOMAIN 8085 8105
FT TRANSSEM 8106 8126
FT DOMAIN 8127 8147
FT TRANSSEM 8148 8168
FT DOMAIN 8169 8189
FT TRANSSEM 8190 8210
FT DOMAIN 8211 8231
FT TRANSSEM 8232 8252
FT DOMAIN 8253 8273
FT TRANSSEM 8274 8294
FT DOMAIN 8295 8315
FT TRANSSEM 8316 8336
FT DOMAIN 8337 8357
FT TRANSSEM 8358 8378
FT DOMAIN 8379 8399
FT TRANSSEM 8400 8420
FT DOMAIN 8421 8441
FT TRANSSEM 8442 8462
FT DOMAIN 8463 8483
FT TRANSSEM 8484 8504
FT DOMAIN 8505 8525
FT TRANSSEM 8526 8546
FT DOMAIN 8547 8567
FT TRANSSEM 8568 8588
FT DOMAIN 8589 8609
FT TRANSSEM 8610 8630
FT DOMAIN 8631 8651
FT TRANSSEM 8652 8672
FT DOMAIN 8673 8693
FT TRANSSEM 8694 8714
FT DOMAIN 8715 8735
FT TRANSSEM 8736 8756
FT DOMAIN 8757 8777
FT TRANSSEM 8778 8798
FT DOMAIN 8799 8819
FT TRANSSEM 8820 8840
FT DOMAIN 8841 8861
FT TRANSSEM 8862 8882
FT DOMAIN 8883 8903
FT TRANSSEM 8904 8924
FT DOMAIN 8925 8945
FT TRANSSEM 8946 8966
FT DOMAIN 8967 8987
FT TRANSSEM 8988 9008
FT DOMAIN 9009 9029
FT TRANSSEM 9030 9050
FT DOMAIN 9051 9071
FT TRANSSEM 9072 9092
FT DOMAIN 9093 9113
FT TRANSSEM 9114 9134
FT DOMAIN 9135 9155
FT TRANSSEM 9156 9176
FT DOMAIN 9177 9197
FT TRANSSEM 9198 9218
FT DOMAIN 9219 9239
FT TRANSSEM 9240 9260
FT DOMAIN 9261 9281
FT TRANSSEM 9282 9302
FT DOMAIN 9303 9323
FT TRANSSEM 9324 9344
FT DOMAIN 9345 9365
FT TRANSSEM 9366 9386
FT DOMAIN 9387 9407
FT TRANSSEM 9408 9428
FT DOMAIN 9429 9449
FT TRANSSEM 9450 9470
FT DOMAIN 9471 9491
FT TRANSSEM 9492 9512
FT DOMAIN 9513 9533
FT TRANSSEM 9534 9554
FT DOMAIN 9555 9575
FT TRANSSEM 9576 9596
FT DOMAIN 9597 9617
FT TRANSSEM 9618 9638
FT DOMAIN 9639 9659
FT TRANSSEM 9660 9680
FT DOMAIN 9681 9701
FT TRANSSEM 9702 9722
FT DOMAIN 9723 9743
FT TRANSSEM 9744 9764
FT DOMAIN 9765 9785
FT TRANSSEM 9786 9806
FT DOMAIN 9807 9827
FT TRANSSEM 9828 9848
FT DOMAIN 9849 9869
FT TRANSSEM 9870 9890
FT DOMAIN 9891 9911
FT TRANSSEM 9912 9932
FT DOMAIN 9933 9953
FT TRANSSEM 9954 9974
FT DOMAIN 9975 9995
FT TRANSSEM 9996 10000

```

```

QY 121 WHNYRQKSMDSKG 134
Db 696 WYNYRQKSMDSKG 709

RESULT 2
ID LGR8_MOUSE STANDARD; PRT; 737 AA.
AC Q91425,
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Relaxin receptor 2 (Leucine-rich repeat-containing G protein-coupled
DE receptor 8) (G protein-coupled receptor affecting testicular descent)
DE (G protein-coupled receptor 106).
DE LGR8 OR GREAT OR GPR106.
GN Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RA MEDLINE=21250990; PubMed=11353515;
RA Overbeek P.A., Gorlov I.P., Sutherland R.W., Houston J.B.,
RA Harrison W.R., Boettger-Tong H.L., Bishop C.E., Agoulnik A.I.;
RA "A transgenic insertion causing cryptorchidism in mice.";
RA Genesis 30:26-35(2001).
FL
CC -I- FUNCTION: Receptor for relaxin. The activity of this receptor is
CC mediated by G proteins leading to stimulation of adenylate cyclase
CC and an increase of cAMP. May also be a receptor for Leydig
CC insulin-like peptide (INSL3) (By similarity).
CC -I- SUBCELLULAR LOCATION: Integral membrane protein.
CC -I- TISSUE SPECIFICITY: Expressed in embryonic and adult gonads of
CC males and females, as well in male gubernaculum. Expressed also
CC in brain. Not detected in kidney, spleen and heart.
CC -I- DISEASE: Defects in LGR8 seems to be a cause of impaired
CC testicular descent (known as cryptorchidism).
CC -I- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -I- SIMILARITY: Contains 1 LDL-receptor class A domain.
CC -I- SIMILARITY: Contains 10 leucine-rich (LRR) repeats.
CC -I- CAUTION: By homology with the human sequence, it is uncertain
CC whether Met-1 is the initiator.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF346501; AAL08943.1; -.
CC MGD; MGI:2153463; Gpr106.
CC GO; GO:0016509; P:protein-hormone receptor activity; IMP.
CC GO; GO:0007189; P:G-protein signaling, adenylate cyclase acti. .; IMP.
CC GO; GO:0008584; P:male gonad development; IMP.
CC InterPro: IPR000276; GPCR_Rhodpsn.
CC InterPro: IPR002172; LDL_receptor_A.
CC InterPro: IPR001611; LRR.
CC InterPro: IPR00372; LRR_Nterm.
CC InterPro: IPR003591; LRR_typ.
CC Pfam; PF00001; 7tm_1; 1.
CC Pfam; PF00057; lcl_recept_a; 1.
CC Pfam; PF00560; LRR; 7.
CC PRINTS; PR00237; GPCRHOOPS.
CC SMART; SM00192; LDLa; 1.
CC SMART; SM00369; LRR_TVP; 2.
CC SMART; SM00013; LRRNT; 1.
CC PROSITE; PS00262; G_PROTEIN_RECP_F1_2; 1.
CC PROSITE; PS01209; LDLRA_1; 1.
CC PROSITE; PS00068; LDLRA_2; 1.

```

KW G-protein coupled receptor; Transmembrane; Glycoprotein; Repeat;  
KW Leucine-rich repeat  
FT DOMAIN 1 399 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 400 420 1 (POTENTIAL).  
FT DOMAIN 421 438 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 439 459 2 (POTENTIAL).  
FT DOMAIN 460 490 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 491 511 3 (POTENTIAL).  
FT DOMAIN 512 520 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 521 541 4 (POTENTIAL).  
FT DOMAIN 542 575 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 576 596 5 (POTENTIAL).  
FT DOMAIN 597 622 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 623 643 6 (POTENTIAL).  
FT DOMAIN 644 664 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 665 737 CYTOPLASMIC (POTENTIAL).  
FT DOMAIN 737 764 LDL-RECEPTOR CLASS A.  
FT REPEAT 119 142 LRR 1.  
FT REPEAT 144 166 LRR 2.  
FT REPEAT 167 190 LRR 3.  
FT REPEAT 192 214 LRR 4.  
FT REPEAT 215 238 LRR 5.  
FT REPEAT 239 262 LRR 6.  
FT REPEAT 263 286 LRR 7.  
FT REPEAT 287 310 LRR 8.  
FT REPEAT 312 334 LRR 9.  
FT REPEAT 335 358 LRR 10.  
FT DISULFID 478 556 BY SIMILARITY.  
FT CARBOHYD 37 37 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 121 121 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 257 257 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 318 318 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 350 350 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 361 361 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 737 AA; 82943 MW; AF60F635EAIACE49 CRC64;

Query Match 49.5%; Score 409; DB 1; Length 737;  
Best Local Similarity 58.9%; Pred. No. 2.3e-30;  
Matches 73; Conservative 27; Mismatches 24; Indels 0; Gaps 0;

QY 4 YSAVAFIGIMAFIIVSYSGMYSVHQSAITAEIRNOVKEMILAKRFFIVTDA 63  
Db 572 YSLGIFGVNLLAFIVISYVYTMFCSTHKTLQATQEVRSHGKEVAVANRFFIVPSDA 631  
QY 64 LCWIPFVAKPLSLOVEIPGTTTSWVIGYSAINALNPILYTLTTRPFKEMIHREWHN 123  
Db 632 ICWIPFVAKPLSLOVEIPGTTTSWVIGYSAINALNPILYTLTTRPFKEMIHREWHN 123  
QY 124 YRQR 127  
Db 692 HRRK 695

RESULT 3  
LGR8\_HUMAN STANDARD; PRT; 754 AA.  
AC Q8WDX0;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DE 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Relaxin receptor 2 (Leucine-rich repeat-containing G protein-coupled receptor 8) (G protein-coupled receptor affecting testicular descent).  
GN LGR8 OR GREAT.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A., AND MUTAGENESIS OF ASP-647.  
RX MEDLINE=21669315; PubMed=11809971;  
RA Hsu S.Y., Nakabayashi K., Nishi S., Kumagai J., Kudo M.,  
RA Sherwood O.D., Hsueh A.J.W.;

\*Activation of orphan receptors by the hormone relaxin.\*;  
Science 295:671-674(2002).  
[2]  
RN SEQUENCE FROM N.A., VARIANT CRYPTORCHIDISM PRO-222, AND VARIANT  
RP VAL-604.  
RX MEDLINE=22206409; PubMed=12217959;  
RA Gorlov I.P., Kanat A., Bogatcheva N.V., Jones E., Lamb D.J.,  
RA Truong A., Bishop C.E., McElreavey K., Agoulnik A.I.;  
RA "Mutations of the GREAT gene cause cryptorchidism.\*";  
Hum. Mol. Genet. 11:2309-2318(2002).  
CC !- FUNCTION: Receptor for relaxin. The activity of this receptor is  
CC mediated by G proteins leading to stimulation of adenylate cyclase  
CC and an increase of cAMP. May also be a receptor for Leydig  
CC insulin-like peptide (INSL3).  
CC !- SUBCELLULAR LOCATION: Integral membrane protein.  
CC !- TISSUE SPECIFICITY: Expressed mainly in the brain kidney, muscle,  
CC testis, thyroid, uterus, peripheral blood cells and bone marrow.  
CC !- DISEASE: Defects in LGR8 are the cause of cryptorchidism (MIM:  
CC 219050). This failure of testicular descent is one of the most  
CC frequent congenital malformations, affecting 1-3% of newborn boys.  
CC Cryptorchidism is associated with increased risk of infertility  
CC and testicular cancer.  
CC !- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.  
CC !- SIMILARITY: Contains 1 LDL-receptor class A domain.  
CC !- SIMILARITY: Contains 10 leucine-rich (LRR) repeats.  
CC !- CAUTION: It is uncertain whether Met-1 or Met-18 is the initiator.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC  
CC EMBL; AF403384; AAL69324.2; -;  
CC EMBL; AF453828; AAL73946.1; -;  
CC MIM; 606655; -;  
CC MIM; 219050; -;  
CC InterPro: IPR000276; GPCR\_Rhodopsn.  
CC InterPro: IPR002172; LDL\_receptor\_A.  
CC InterPro: IPR001611; LRR.  
CC InterPro: IPR000372; LRR\_Nterm.  
CC InterPro: IPR003591; LRR\_typ.  
CC Pfam: PF00001; 7tm.1; 1  
CC Pfam: PF00057; ldl\_recept\_a; 1.  
CC Pfam: PF00560; LRR; 7.  
CC PRINTS: PR00237; GPCRHOPOPSN.  
CC SMART; SM00132; LDLa; 1.  
CC SMART; SM00369; LRR\_TYP; 3.  
CC SMART; SM00013; LRRT; 1.  
CC PROSITE: PS00237; G\_PROTEIN\_RECP\_F1\_1; FALSE\_NEG.  
CC PROSITE: PS50262; G\_PROTEIN\_RECP\_F1\_2; 1.  
CC PROSITE: PS01209; LDLRA\_1; 1.  
CC PROSITE: PS50068; LDLRA\_2; 1.  
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Repeat;  
KW Leucine-rich repeat; Polymorphism; Disease mutation.  
FT DOMAIN 1 416 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 417 437 1 (POTENTIAL).  
FT DOMAIN 438 455 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 456 476 2 (POTENTIAL).  
FT DOMAIN 477 507 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 508 528 3 (POTENTIAL).  
FT DOMAIN 529 537 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 538 558 4 (POTENTIAL).  
FT DOMAIN 559 592 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 593 613 5 (POTENTIAL).  
FT DOMAIN 614 639 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 640 660 6 (POTENTIAL).  
FT DOMAIN 661 673 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 674 694 7 (POTENTIAL).  
FT DOMAIN 695 754 CYTOPLASMIC (POTENTIAL).  
FT DOMAIN 754 81 LDL-RECEPTOR CLASS A.

```

FT REPEAT 136 159 LRR 1.
FT REPEAT 161 183 LRR 2.
FT REPEAT 184 207 LRR 3.
FT REPEAT 209 231 LRR 4.
FT REPEAT 233 255 LRR 5.
FT REPEAT 256 279 LRR 6.
FT REPEAT 281 303 LRR 7.
FT REPEAT 304 327 LRR 8.
FT REPEAT 329 351 LRR 9.
FT REPEAT 352 375 LRR 10.
FT DISULFID 495 573 BY SIMILARITY.
FT CARBOHYD 54 54 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 138 138 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 274 274 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 335 335 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 378 378 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARIANT 222 222 T -> P (in cryptorchidism; functionally inactive).
FT FTID=VAR_015386.
FT VARIANT 604 604 I -> V.
FT MUTAGEN 647 647 D->Y: LEADS TO CONSTITUTIVE INCREASE OF BASAL CAMP.
SQ SEQUENCE 754 AA; 86452 MW; 2088ECD204C6A6C5 CRC64;

Query Match 47.5%; Score 392.5; DB 1; Length 754;
Best Local Similarity 59.1%; Pred. No. 8e-23;
Matches 75; Conservative 25; Mismatches 26; Indels 1; Gaps 1;

Qy 4 YSAVAFGLINLAFLIVSYGSMFYSHOSAITATATIRNOVKEMILAKSEFFIVETDA 63
Db 589 YSLGIFGLNLAFLIVSYITMFCISQKTALQTTEVRCFGREAVANREFFIVESDA 648
Qy 64 LCWIPFYAKPLSLQVEIPGTHITSMVYIGYSAINSLNPILYTLTRPFKEMIHREWHN 123
Db 649 ICWIPFVVKLSLFRVLEIPDTMTSWIVTFPLVNSALNPILYTLTRPFKDKLQLLKH 708
Qy 124 YRQKSM 130
Db 709 H-ORCSI 714

RESULT 4
GPCR_LYMT STANDARD; PRT; 1115 AA.
AC P46023;
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE G-protein coupled receptor GRL101 precursor.
OS Lymnaea stagnalis (Great pond snail).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Basommatophora;
OC Lymnaeidae; Lymnaeidae; Lymnaeae.
OX NCBI_Taxid=6523;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=CNS;
RX MEDLINE=94255418; PubMed=8197140;
RA Tensen C.P., van Kesteren E.R., Planta R.J., Cox K.J.A., Burke J.F.,
RA van Heerikhuizen H., Vreugdenhil E.;
RT "A G protein-coupled receptor with low density lipoprotein-binding
RT motifs suggests a role for lipoproteins in G-linked signal
transduction.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:4816-4820(1994).
CC -1- FUNCTION: MIGHT DIRECTLY TRANSDUCE SIGNALS CARRIED BY LARGE
CC EXTRACELLULAR (LIPO)PROTEIN(COMPLEXES) INTO NEURONAL EVENTS.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN A SMALL NUMBER OF
CC NEURONS WITHIN THE CENTRAL NERVOUS SYSTEM AND TO A LESSER EXTENT
CC IN THE HEART.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -1- SIMILARITY: Contains 12 LDL-receptor class A domains.
CC -1- SIMILARITY: Contains 7 leucine-rich (LRR) repeats.

```

```

CC -----PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z23104; CAA80651.1; -
CC FIR; S40241; S40241.
CC HSP; Q07954; ICR8.
CC InterPro; IPR000276; GPCR_Rhodpsn.
CC InterPro; IPR002172; LDL_receptor_A.
CC InterPro; IPR001611; LRR.
CC InterPro; IPR000372; LRR_Nterm.
CC InterPro; IPR003591; LRR_typ.
CC Pfam; PF00001; Tm1_1; 1.
CC Pfam; PF00057; ldl_recept_a; 11.
CC Pfam; PF00560; LRR; 6.
CC Pfam; PF01462; LRRNT; 1.
CC PRINTS; PR00237; GPCRRHODOPSN.
CC PRINTS; PR00261; LDLRECEPTOR.
CC SMART; SM00192; Lda; 12.
CC SMART; SM00369; LRR_TYP; 1.
CC SMART; SM00013; LRRNT; 1.
CC PROSITE; PS00237; G_PROTEIN_RECP_FT_1; FALSE_NEG.
CC PROSITE; PS0262; G_PROTEIN_RECP_FT_2; 1.
CC PROSITE; PS01209; LDLRA_1; 6.
CC PROSITE; PS00068; LDLRA_2; 11.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Repeat;
KW Leucine-rich repeat; Signal.
FT SIGNAL 1 24
FT CHAIN 25 1115 POTENTIAL.
FT DOMAIN 25 767 G-PROTEIN COUPLED RECEPTOR GRL101.
FT TRANSMEM 768 788 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 789 801 1 (POTENTIAL).
FT TRANSMEM 802 822 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 823 857 2 (POTENTIAL).
FT TRANSMEM 858 878 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 879 887 3 (POTENTIAL).
FT TRANSMEM 888 908 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 909 941 4 (POTENTIAL).
FT TRANSMEM 942 962 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 963 988 5 (POTENTIAL).
FT TRANSMEM 989 1009 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 1010 1017 6 (POTENTIAL).
FT TRANSMEM 1018 1038 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 1039 1115 7 (POTENTIAL).
FT TRANSMEM 1039 1115 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 32 523 12 X 40 AA APPROXIMATE TANDEM REPEATS
FT TRANSMEM 36 79 SIMILAR TO THE LDL-RECEPTOR CLASS A.
FT DOMAIN 77 115 LDL-RECEPTOR CLASS A 1.
FT DOMAIN 116 155 LDL-RECEPTOR CLASS A 2.
FT DOMAIN 156 196 LDL-RECEPTOR CLASS A 3.
FT DOMAIN 195 232 LDL-RECEPTOR CLASS A 4.
FT DOMAIN 231 269 LDL-RECEPTOR CLASS A 5.
FT DOMAIN 272 318 LDL-RECEPTOR CLASS A 6.
FT DOMAIN 320 363 LDL-RECEPTOR CLASS A 7.
FT DOMAIN 365 403 LDL-RECEPTOR CLASS A 8.
FT DOMAIN 404 442 LDL-RECEPTOR CLASS A 9.
FT DOMAIN 444 485 LDL-RECEPTOR CLASS A 10.
FT DOMAIN 486 525 LDL-RECEPTOR CLASS A 11.
FT REPEAT 557 579 LDL-RECEPTOR CLASS A 12.
FT REPEAT 582 603 LRR 1.
FT REPEAT 607 629 LRR 2.
FT REPEAT 630 653 LRR 3.
FT REPEAT 655 677 LRR 4.
FT REPEAT 678 701 LRR 5.
FT REPEAT 703 725 LRR 6.
FT DISULFID 38 53 LRR 7.
FT DISULFID 46 56 BY SIMILARITY.
FT DISULFID 60 77 BY SIMILARITY.

```

```

RN 1 AQTYSVAIFGINLAAPFIIVFSYMSFVSHQSAITATEIRNOVKKEMILAKRFFFIYV 60
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RA Yadatie F., Male R.;
RT receptor (TSHR) like cDNA;
RL Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC EMBL; AF406603; AAK98600.1;
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm.1.1.
DR PRINTS; PR00237; GPCRHHODOPS.
DR PROSITE; PS00237; G_PROTEIN_RECF_F1_1.
DR PROSITE; PS0262; G_PROTEIN_RECF_F1_2; 1.
KW G-protein coupled receptor; Receptor; Transmembrane.
FT NON_TER 1
SQ SEQUENCE 307 AA; 34289 MW; 0E3B8032DAF68470 CRC64;

Query Match 16.7%; Score 138; DB 13; Length 307;
Best Local Similarity 28.9%; Pred. No. 4.1e-06;
Matches 35; Conservative 25; Mismatches 41; Indels 20; Gaps 4;

QY 1 AQTYSVAIFGINLAAPFIIVFSYMSFVSHQSAITATEIRNOVKKEMILAKRFFFIYV 60
Db 115 AQTYSVLI-LNLAFLVICYTKIYVHNPHYRSG-----SKDNIARKMAVLIF 167
QY 61 TDALCWIPF-----FVAKPLSLLOVEIPGTITSVWVIGYSAINSLNPILYTLTPRF 113
Db 168 TDFLCMAPIFSYMSFVMDREL-----ITVNSKILLVFLPLNSCANFLYAIPTKAF 221
QY 114 K 114
Db 222 R 222

RESULT 9
QY1949 PRELIMINARY; PRT; 814 AA.
AC Q91949;
DT 01-OCT-2000 (TREMREL. 15, Created)
DT 01-OCT-2000 (TREMREL. 15, Last sequence update)
DE 01-MAR-2003 (TREMREL. 23, Last annotation update)
DE Thyrotropin receptor A.
GN STSH-RA.
OS Oncomyrhynchus rhodurus (amago).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncomyrhynchus.
OX NCBI_TaxID=41164;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20462948; Pubmed=11006115;
RA Oba Y., Hirai T., Yoshiura Y., Kobayashi T., Nagahama Y.;
RT "Cloning, Functional Characterization, and Expression of Thyrotropin
Receptors in the Thyroid of Amago Salmon (Oncomyrhynchus rhodurus).";
RL Biochem. Biophys. Res. Commun. 276:258-263(2000).
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC EMBL; AB030954; BAB07800.1;
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR EMBL; AB030954; BAB07800.1;
DR InterPro; IPR001611; LRR.
DR Pfam; PF00001; 7tm.1.1.
DR Pfam; PF00560; LRR.1.
DR PRINTS; PR00237; GPCRHHODOPS.
DR PROSITE; PS00237; G_PROTEIN_RECF_F1_1; 1.
DR PROSITE; PS0262; G_PROTEIN_RECF_F1_2; 1.
KW G-protein coupled receptor; Receptor; Transmembrane.
SQ SEQUENCE 814 AA; 91091 MW; 1FC9A2AE8D6C911E CRC64;

Query Match 16.2%; Score 134; DB 13; Length 814;
Best Local Similarity 26.8%; Pred. No. 3e-05;
Matches 44; Conservative 32; Mismatches 60; Indels 28; Gaps 6;

```

```

QY 1 AQTYSVAIFGINLAAPFIIVFSYMSFVSHQSAITATEIRNOVKKEMILAKRFFFIYV 60
Db 624 AQTYSVLI-LNLAFLVICYTKIYVHNPHYRSG-----SKDNIARKMAVLIF 676
QY 61 TDALCWIPF-----FVAKPLSLLOVEIPGTITSVWVIGYSAINSLNPILYTLTPRF 114
Db 677 TDFLCMAPIFSYMSFVMDREL-----ITVNSKILLVFLPLNSCANFLYAIPTKAF 736
QY 115 -----EMTHRWNYRQKSKMDSKGRKMHLSHSSGKGCRCRR 153
Db 737 LLSKVGLCORRAQLFR-GQTVSSKG-----SSGVCQGR 770

RESULT 10
QYDE63 PRELIMINARY; PRT; 424 AA.
AC Q9DE63;
DT 01-MAR-2001 (TREMREL. 16, Created)
DT 01-MAR-2001 (TREMREL. 16, Last sequence update)
DT 01-DEC-2001 (TREMREL. 19, Last annotation update)
DE GnRH receptor-3 (Type III GnRH).
GN BGNRRH-3.
OS Rana catesbeiana (Bull frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranioidea; Ranidae; Rana.
OX NCBI_TaxID=8400;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21065155; Pubmed=11120886;
RA Wang L., Bogerd J., Choi H.S., Seong J.Y., Soh J.M., Chun S.Y.,
RA Blumenrohr M., Troskie B.E., Millar R.P., Yu W.H., McCann S.M.,
RA Xwon H.B.;
RT "Three distinct types of GnRH receptor characterized in the
bullfrog.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:361-366(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21407809; Pubmed=11517181;
RA Wang L., Oh D.Y., Bogerd J., Choi H.S., Ahn R.S., Seong J.Y.,
RA Xwon H.B.;
RT "Inhibitory activity of alternative splice variants of the bullfrog
GnRH receptor-3 on wild-type receptor signaling.";
RL Endocrinology 142:4015-4025(2001).
DR EMBL; AF144062; AAG42574.1;
DR EMBL; AF224277; AAL11631.1;
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm.1.1.
DR PRINTS; PR00237; GPCRHHODOPS.
DR PROSITE; PS0262; G_PROTEIN_RECF_F1_2; 1.
KW Receptor.
SQ SEQUENCE 424 AA; 47613 MW; E771F7403AB44AD6 CRC64;

Query Match 16.2%; Score 133.5; DB 13; Length 424;
Best Local Similarity 22.6%; Pred. No. 1.7e-05;
Matches 44; Conservative 33; Mismatches 73; Indels 45; Gaps 6;

QY 3 IYSVAIFGINLAAPFIIVFSYMSFVSHQSAITATEIRNOVKKEMILAKRFFFIYV 52
Db 224 LYNNFTFCFLPLLLIWFVGYGRILVLSRKKKAQVSSREYNLRSSNNIPRMRTE 283
QY 53 KRFFFTVETDALCWIPFVA-----KPLSLLOVEIPGTITSVWVIGYSAINSLNPILY 106
Db 284 KMSLVITLTFVTCWTPYLLGIWTFSPENITSRKVPPLSL-SHILFLGLFTCLDPLII 342
QY 107 TLTPRFKEMHFRWNYRQKSKMDSKGRKMHLSHSSGKGCRCRR 145
Db 343 GLFTIHRFRVRCATGKADATSLTGTSPFIATAVPLKRSAGSGSGCKFDLEV 402
QY 146 -----GKQGR 153
Db 403 TGVLHSGKCEHRR 417

```

5.

```

Db      641 TDFLCMAPISFYANSAVLDRLPITVNSKILLVLPYPLNSCANPFLYAIFTKAPR 695

RESULT 14
Q98TF4      PRELIMINARY;      PRT;      778 AA.
AC  Q98TF4; 2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DI 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Thyrotropin receptor.
GN TSH-R.
OS Oreochromis niloticus (Nile tilapia) (Tilapia nilotica).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Notoleleostei;
OC Acanthomorphia; Acanthopterygii; Percomorpha; Perciformes; Labroidae;
OC Cichlidae; Oreochromis.
OX NCBI_TaxID=8128;
RN [1]
RP SEQUENCE FROM N.A.
RA Oba Y.; Nagahama Y.;
RT "tilapia thyrotropin receptor.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR EMBL; AB047390; BAB39132.1; -.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR InterPro; IPR001611; LRR.
DR Pfam; PF00001; 7tm_1; 1.
DR Pfam; PF00560; LRR; 2.
DR PRINTS; PR00237; GPCRHHODOPSN.
DR PROSITE; PS00237; G-PROTEIN_RECEP_FL_1; 1.
DR PROSITE; PS50262; G-PROTEIN_RECEP_FL_2; 1.
KW G-protein coupled receptor; Receptor; Transmembrane.
SQ SEQUENCE 778 AA; 87126 MW; 618A314295880A9E CRC64;

Query Match      15.8%; Score 130.5; DB 13; Length 778;
Best Local Similarity 28.0%; Pred. No. 6.5e-05;
Matches 49; Conservative 30; Mismatches 67; Indels 29; Gaps 7;

QY      1 AQIYSVAIFLGINLAAFIIIVFSYGSMTFVSHQSAITATETIRNQVKKEMILAKRFFPIVF 60
      ||||| : : : : : || : : : : : : : : : : : : : : : : : : : : : : : :
Db      587 AQIYIMSV-LVNLIAFFIACFYKIYCTVHPHY-----HSGSKDTNIAKRMVAVLF 639

QY      61 TDALCWIP-FYAKPLSLQVEIPGTTISWVWIGYSAINSLNPLYLTLTPFKEMLHR 119
      || || || || || || || || || || || || || || || || || || || || ||
Db      640 TDFLCMAPISFYANSAVLDRLPITVNSKILLVLPYPLNSCANPFLYAIFTKAPR 699

QY      120 FWHNY---RQR-----KSMDSKG-----IRKMLHSHSGKGC-----HCRR 153
      : : : : : || : : : : : || : : : : : || : : : : : || : : : : :
Db      700 LLSKVLGQCQRAQLFRGQTVSSKSGSGTSQIRRDKDKSRKGGSGIQBEVPIHLKR 754

RESULT 15
Q9EN18      PRELIMINARY;      PRT;      1050 AA.
AC  Q9EN18;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DI 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DI 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Leucine-rich repeat-containing G protein-coupled receptor 2.
GN RK OR CG8930.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Nishi S.; Hsu S.Y.; Zell K.; Hsueh A.J.;
RT "Characterization of two fly LGR (leucine-rich repeat-containing G
protein-coupled receptor) proteins homologous to vertebrate
glycoprotein hormone receptors: constitutively activation of wild type

```

```

RT fly LGR1 but not LGR2 in transfected mammalian cells.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF274591; AAK00808.1; -.
DR FlyBase; FBgn0003255; rk.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR003591; LRR_tyr.
DR Pfam; PF00001; 7tm_1; 1.
DR Pfam; PF00560; LRR; 10.
DR PRINTS; PR00237; GPCRHHODOPSN.
DR SMART; SM00369; LRR_TYP; 4.
DR PROSITE; PS50262; G-PROTEIN_RECEP_FL_2; 1.
DR PROSITE; PS50506; LRR_TYPICAL; 2.
KW Receptor.
SQ SEQUENCE 1050 AA; 117707 MW; 35D71260A8B4BF99 CRC64;

Query Match      15.7%; Score 130; DB 5; Length 1050;
Best Local Similarity 30.9%; Pred. No. 0.0001;
Matches 38; Conservative 21; Mismatches 42; Indels 22; Gaps 4;

QY      1 AQIYSVAIFLGINLAAFIIIVFSYGSMTFVSHQSAITATETIRNQVKKEMILAKRFFPIVF 60
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db      704 ASLTIVSIFMGCAFLTLMGCTLMYWAIRGSAWNT-----NDSRIAKRMALLVF 756

QY      61 TDALCWIP-----IFVAKPLSLQVEIPGTTISWVWIGYSAINSLNPLYLTLTPR 112
      || || || || || || || || || || || || || || || || || || || || ||
Db      757 TDFLCSPIAEFSITAIFGLQLISLEQAKI---FTVFVL-----PLASCONPFLYAINTKQ 809

QY      113 FKE 115
      ||:
Db      810 FKK 812

Search completed: September 25, 2003, 15:00:38
Job time : 38 secs

```



```

RX MEDLINE-92149579; PubMed-1738373;
RA Heckert L.L., Daley I.J., Griswold M.D.;
RT "Structural organization of the follicle-stimulating hormone receptor
   gene.";
RL Mol. Endocrinol. 6:70-80(1992).
CC -!- FUNCTION: RECEPTOR FOR FOLLICLE STIMULATING HORMONE. THE ACTIVITY
CC OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE
CC ADENYLATE CYCLASE.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: SEROTOL CELLS AND OVARIAN GRANULOSA CELLS.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC FSH/LSH/TSH SUBFAMILY.
CC -!- SIMILARITY: Contains 6 leucine-rich (LRR) repeats.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; L02842; AAA41175.1; -.
CC PIR; A34548; A34548.
CC HSSP; P23945; 1XUN.
CC InterPro; IPR000276; GPCR_Rhodpsn.
CC InterPro; IPR001611; LRR_Nterm.
CC Pfam; PF000560; LRR; 3.
CC Pfam; PF01462; LRRNT; 1.
CC PRINTS; PR00237; GPCRHHODPSN.
CC SMART; SM00013; LRRNT; 1.
CC PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
CC PROSITE; PS02626; G_PROTEIN_RECEP_F1_2; 1.
CC G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
CC Phosphorylation; Repeat; Leucine-rich repeat.
CC SIGNAL 1 17
CC CHAIN 18 692 FOLLICLE STIMULATING HORMONE RECEPTOR.
CC DOMAIN 18 365 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 366 386 1.(POTENTIAL).
CC DOMAIN 387 397 CYTOPLASMIC (POTENTIAL).
CC TRANSMEM 398 420 2 (POTENTIAL).
CC DOMAIN 421 442 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 443 464 3 (POTENTIAL).
CC DOMAIN 465 484 CYTOPLASMIC (POTENTIAL).
CC TRANSMEM 485 507 4 (POTENTIAL).
CC DOMAIN 508 527 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 528 549 5 (POTENTIAL).
CC DOMAIN 550 572 CYTOPLASMIC (POTENTIAL).
CC TRANSMEM 573 596 6 (POTENTIAL).
CC DOMAIN 597 607 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 608 629 7 (POTENTIAL).
CC DOMAIN 630 692 CYTOPLASMIC (POTENTIAL).
CC REPEAT 44 68 LRR 1.
CC REPEAT 69 93 LRR 2.
CC REPEAT 119 143 LRR 3.
CC REPEAT 170 192 LRR 4.
CC REPEAT 193 216 LRR 5.
CC REPEAT 218 240 LRR 6.
CC REPEAT 441 516 BY SIMILARITY.
CC DISULFID 441 516 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 191 191 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 199 199 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 293 293 N-LINKED (GLCNAC. .) (POTENTIAL).
CC SEQUENCE 692 AA; 77681 MW; 267EA78C7CFD8E6 CRC64;
Query Match 16.98; Score 139.5; DB 1; Length 692;
Best Local Similarity 27.38; Pred. No. 1.8e-05;
Matches 47; Conservative 30; Mismatches 60; Indels 35; Gaps 6;
OY 1 AQIVSVAIFGLGNAAPFIIVTSYSGMFSYQSQAITATEIRNQVKEMILAKRFFFIIVF 60
DB 526 SGLYVMAL-LVLNVLAIFVVICGYTHIVLRNPTIVSS-----SSDTKIAKEMALIF 578

```

```

OY 61 TDALCWIPF-FVAKPLSLQVLEIPGTTISWVIGVSAINSGALMPLYLTTRPFKE---- 115
DB 579 TDFLCMAPISTFAISAKLVPLITVLFYFIPINSCANPFLYAITFNFRDFFI 638
OY 116 MIHRE-----WENYRQRKSMDSKGIRKH-----MLHSS 144
DB 639 LUSKFGCIEMOAIQYRTTSATNHFARKSCSAPAVNSVLYPLNHS 690

RESULT 10
FSHR_BOVIN
ID FSHR_BOVIN STANDARD; PRT; 695 AA.
AC P33376;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Follicle stimulating hormone receptor precursor (FSH-R) (Follitropin
DE receptor).
GN FSHR.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Holstein; TISSUE=Ovary. and Testis;
RX MEDLINE=95127199; PubMed=7826612;
RA Houde A., Lambert A., Saumande J., Silversides D.W., Lussier J.G.;
RT "Structure of the bovine follicle-stimulating hormone receptor
RT complementary DNA and expression in bovine tissues.";
RL Mol. Reprod. Dev. 39:127-135(1994).
CC -!- FUNCTION: RECEPTOR FOR FOLLICLE STIMULATING HORMONE. THE ACTIVITY
CC OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE
CC ADENYLATE CYCLASE.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC FSH/LSH/TSH SUBFAMILY.
CC -!- SIMILARITY: Contains 6 leucine-rich (LRR) repeats.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; L22319; AAC37324.1; -.
CC PIR; I45896; I45896.
CC HSSP; P23945; 1XUN.
CC InterPro; IPR000276; GPCR_Rhodpsn.
CC InterPro; IPR001611; LRR_Nterm.
CC Pfam; PF00001; 7tm_1; 1.
CC Pfam; PF00560; LRR; 5.
CC Pfam; PF01462; LRRNT; 1.
CC PRINTS; PR00237; GPCRHHODPSN.
CC SMART; SM00013; LRRNT; 1.
CC PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
CC PROSITE; PS02626; G_PROTEIN_RECEP_F1_2; 1.
CC G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
CC Phosphorylation; Repeat; Leucine-rich repeat.
CC SIGNAL 1 17 POTENTIAL.
CC CHAIN 18 695 FOLLICLE STIMULATING HORMONE RECEPTOR.
CC DOMAIN 18 366 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 367 387 1 (POTENTIAL).
CC DOMAIN 388 398 CYTOPLASMIC (POTENTIAL).
CC TRANSMEM 399 421 2 (POTENTIAL).
CC DOMAIN 422 443 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 444 465 3 (POTENTIAL).
CC DOMAIN 466 485 CYTOPLASMIC (POTENTIAL).

```

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----CC

EMBL: U80673; AAB53698.1; -

DR HSP: P22888; ILUT.

DR InterPro: IPR000276; GPCR\_Rhodpsn.

DR InterPro: IPR001611; LRR\_Nterm.

DR InterPro: IPR000372; LRR\_Nterm.

DR Pfam: PF00001; 7tm\_1; 1.

DR Pfam: PF00560; LRR; 2.

DR PRINTS: PRO0237; GPCRHHODPSN.

DR SMART: SM00013; LRRT; 1.

DR PROSITE: PS00237; G\_PROTEIN\_RECEP\_F1\_1; 1.

DR PROSITE: PS02662; G\_PROTEIN\_RECEP\_F1\_2; 1.

KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;

KW Phosphorylation; Repeat; Leucine-rich repeat.

FT SIGNAL 1 29

FT CHAIN 30 676

FT LUTROPIN-CHORIOGONADOTROPIC HORMONE

FT RECEPTOR.

FT DOMAIN 30 340

FT TRANSNEM 341 362

FT DOMAIN 363 372

FT TRANSNEM 373 393

FT DOMAIN 394 416

FT TRANSNEM 417 439

FT DOMAIN 440 459

FT TRANSNEM 460 482

FT DOMAIN 483 502

FT TRANSNEM 503 526

FT DOMAIN 527 547

FT TRANSNEM 548 571

FT DOMAIN 572 582

FT TRANSNEM 583 604

FT DOMAIN 605 676

FT REPEAT 52 75

FT REPEAT 126 151

FT REPEAT 153 175

FT REPEAT 176 200

FT REPEAT 225 248

FT DISULFID 416 491

FT CARBOHYD 103 108

FT CARBOHYD 178 178

FT CARBOHYD 199 199

FT SEQUENCE 676 AA; 75677 MW; F8018227641654E1 CRC64;

Query Match 16.8%; Score 139; DB 1; Length 676;

Best Local Similarity 31.3%; Pred. No. 2e-05;

Matches 36; Conservative 24; Mismatches 47; Indels 8; Gaps 3;

QY 1 A QIYVAIEFAGINLAFTIIIVFSYGNFYSVHSAITAEIRNOVKEMILAKRFFFIYF 60

DB 501 SQIYIYTLI-LNVAFIICACIKIYFAVNPMLAT-----NKDTKIANKMAILIF 553

QY 61 TDLACWIPI-FVAKPLSLQLQVEIDGITSWVIGYSAINSLNPILYTLTRPEK 114

DB 554 TDTFCAPISFEFFAISAAKMPLTITVNSKVLIVLFYIPINSCANPFLYAIFTKTRF 608

RESULT 12

FSRR\_PIG

ID FSRR\_PIG STANDARD; PRT; 695 AA.

AC P49059; O77514;

CD 01-FEB-1996 (Rel. 33, Created)

DT 15-DEC-1998 (Rel. 37, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Follicle stimulating hormone receptor precursor (FSH-R) (Follitropin receptor).

DE DE

GN FSHR.

OS Sus scrofa (Pig).

OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Cetartiodactyla; Suina; Suidae; Sus.

OC NCBI TaxID: 9823;

[illegible]

QY 116 MHRF 120  
:: ||  
Db 641 LLSRF 645

RESULT 14  
LSHR RAT  
ID LSHR RAT STANDARD; PRT: 700 RA.  
AC P16235; l70646; Q63807; Q63808; Q63809;  
DT 01-APR-1990 (Rel. 14, Created)  
DT 01-APR-1990 (Rel. 14, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Lutropin-choriogonadotropic hormone receptor precursor (1H/CG-R)  
DE (LSH-R) (luteinizing hormone receptor).  
GN LHCG  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus;  
OC NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=9332512; PubMed=2502842;  
RA McFarland K.C., Sprengel R., Phillips H.S., Koehler M.,  
RA Rosenblit N., Nikolics K., Segaloff D.L., Seeburg P.H.;  
RT "Lutropin-choriogonadotropin receptor: an unusual member of the G  
RT protein-coupled receptor family.";  
RL Science 245:494-499(1989).  
RN [2]  
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.  
RC STRAIN=Sprague-Dawley; TISSUE=Ovary;  
RX MEDLINE=92347604; PubMed=1353463;  
RA Aatsinki J.T., Pietila E.M., Lakkakorpi J.T., Rajaniemi H.J.;  
RT "Expression of the LH/CG receptor gene in rat ovarian tissue is  
RT regulated by an extensive alternative splicing of the primary  
RT transcript.";  
RL Mol. Cell. Endocrinol. 84:127-135(1992).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91209270; PubMed=2019252;  
RA Xoo Y.B., Slaughter R.G., Ji T.H.;  
RT "Structure of the luteinizing hormone receptor gene and multiple  
RT exons of the coding sequence.";  
RL Endocrinology 128:2297-2308(1991).  
RN [4]  
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.  
RX MEDLINE=91006819; PubMed=1975554;  
RA Bernard M.P., Myers R.V., Moyle W.R.;  
RT "Cloning of rat lutropin (LH) receptor analogs lacking the soybean  
RT lectin domain.";  
RL Mol. Cell. Endocrinol. 71:R19-R23(1990).  
RN [5]  
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.  
RX MEDLINE=91126285; PubMed=2281186;  
RA Segaloff D.L., Sprengel R., Nikolics K., Ascoli M.;  
RT "Structure of the lutropin/choriogonadotropin receptor.";  
RL Recent Prog. Horm. Res. 46:261-303(1990).  
RN [6]  
RP SEQUENCE OF 295-700 FROM N.A.  
RX MEDLINE=91060531; PubMed=2174034;  
RA Tsai-Morris C.H., Buczek E., Wang W., Dufau M.L.;  
RT "Intronic nature of the rat luteinizing hormone receptor gene defines  
RT a soluble receptor subpecies with hormone binding activity.";  
RL J. Biol. Chem. 265:19385-19388(1990).  
RN [7]  
RP SEQUENCE OF 27-37.  
RX MEDLINE=89174723; PubMed=2925659;  
RA Roche P.C., Ryan R.J.;  
RT "Purification, characterization, and amino-terminal sequence of rat  
RT ovarian receptor for luteinizing hormone/human chorionadotropin.";  
RL J. Biol. Chem. 264:4636-4641(1989).  
RN [8]  
RP MUTAGENESIS.  
RN RP

```

RX MEDLINE=911332007; PubMed=1714448;
RA Ji I., Ji T.H.;
RT "Asp383 in the second transmembrane domain of the lutropin receptor
RT is important for high affinity hormone binding and cAMP production.";
RL J. Biol. Chem. 266:14953-14957(1991).
CC !- FUNCTION: RECEPTOR FOR LUTROPIN-CHORIOGNADOTROPIC HORMONE. THE
CC ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE
CC ADENYLATE CYCLASE.
CC !- SUBCELLULAR LOCATION: Integral membrane protein. Some isoforms may
CC be secreted.
CC !- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=11;
CC Comment=additional isoforms seem to exist;
CC Name=1;
CC IsoId=P16235-1; Sequence=Displayed;
CC Name=1759;
CC IsoId=P16235-2; Sequence=VSP_001969, VSP_001977, VSP_001978;
CC Name=1834;
CC IsoId=P16235-3; Sequence=VSP_001977, VSP_001978;
CC Name=1950;
CC IsoId=P16235-4; Sequence=VSP_001968;
CC Name=2075;
CC IsoId=P16235-5; Sequence=VSP_001971, VSP_001973;
CC Name=C1;
CC IsoId=P16235-6; Sequence=VSP_001975, VSP_001976;
CC Name=C2;
CC IsoId=P16235-7; Sequence=VSP_001970;
CC Name=RA2;
CC IsoId=P16235-8; Sequence=VSP_001972;
CC Name=EB;
CC IsoId=P16235-9; Sequence=VSP_001972, VSP_001977, VSP_001978;
CC Name=B1;
CC IsoId=P16235-10; Sequence=VSP_001972, VSP_001974, VSP_001979;
CC Name=B3;
CC IsoId=P16235-11; Sequence=VSP_001974, VSP_001979;
CC !- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC FSH/LSH/TSH SUBFAMILY.
CC !- SIMILARITY: Contains 7 leucine-rich (LRR) repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M26199; AAA41528.1; -.
DR EMBL; M61212; AAA41527.1; -.
DR EMBL; M61211; AAA41527.1; JOINED.
DR EMBL; M40803; AAB22680.1; -.
DR EMBL; S40787; AAB22680.1; JOINED.
DR EMBL; S40903; AAB22680.1; JOINED.
DR EMBL; S40904; AAB22680.1; JOINED.
DR EMBL; S40905; AAB22680.1; JOINED.
DR EMBL; S40907; AAB22680.1; JOINED.
DR EMBL; S40909; AAB22680.1; JOINED.
DR EMBL; S40918; AAB22680.1; JOINED.
DR EMBL; S40920; AAB22680.1; JOINED.
DR EMBL; S40795; AAB22680.1; JOINED.
DR EMBL; S40798; AAB22680.1; JOINED.
DR EMBL; S40799; AAB22681.1; -.
DR EMBL; S40787; AAB22681.1; JOINED.
DR EMBL; S40903; AAB22681.1; JOINED.
DR EMBL; S40904; AAB22681.1; JOINED.
DR EMBL; S40905; AAB22681.1; JOINED.
DR EMBL; S40907; AAB22681.1; JOINED.
DR EMBL; S40909; AAB22681.1; JOINED.
DR EMBL; S40918; AAB22681.1; JOINED.
DR EMBL; S40920; AAB22681.1; JOINED.
DR EMBL; S40803; AAB22682.2; -.
DR EMBL; S40787; AAB22682.2; JOINED.
DR EMBL; S40903; AAB22682.2; JOINED.

```

```

DR EMBL; S40907; AAB22682.2; JOINED.
DR EMBL; S40909; AAB22682.2; JOINED.
DR EMBL; S40918; AAB22682.2; JOINED.
DR EMBL; S40920; AAB22682.2; JOINED.
DR EMBL; S40795; AAB22682.2; JOINED.
DR EMBL; S40798; AAB22682.2; JOINED.
DR EMBL; S40803; AAB22683.1; -.
DR EMBL; S40787; AAB22683.1; JOINED.
DR EMBL; S40903; AAB22683.1; JOINED.
DR EMBL; S40904; AAB22683.1; JOINED.
DR EMBL; S40905; AAB22683.1; JOINED.
DR EMBL; S40907; AAB22683.1; JOINED.
DR EMBL; S40909; AAB22683.1; JOINED.
DR EMBL; S40918; AAB22683.1; JOINED.
DR EMBL; S40920; AAB22683.1; JOINED.
DR EMBL; S40795; AAB22683.1; JOINED.
DR EMBL; S40798; AAB22683.1; JOINED.
DR EMBL; S40803; AAB22684.2; -.
DR EMBL; S40787; AAB22684.2; JOINED.
DR EMBL; S40903; AAB22684.2; JOINED.
DR EMBL; S40904; AAB22684.2; JOINED.
DR EMBL; S40905; AAB22684.2; JOINED.
DR EMBL; S40909; AAB22684.2; JOINED.
DR EMBL; S40918; AAB22684.2; JOINED.
DR EMBL; S40920; AAB22684.2; JOINED.
DR EMBL; S40795; AAB22684.2; JOINED.
DR EMBL; S40798; AAB22684.2; JOINED.
DR EMBL; M68928; AAA41529.1; -.
DR EMBL; M68917; AAA41529.1; JOINED.
DR EMBL; M68918; AAA41529.1; JOINED.
DR EMBL; M68919; AAA41529.1; JOINED.
DR EMBL; M68920; AAA41529.1; JOINED.
DR EMBL; M68921; AAA41529.1; JOINED.
DR EMBL; M68922; AAA41529.1; JOINED.
DR EMBL; M68923; AAA41529.1; JOINED.
DR EMBL; M68925; AAA41529.1; JOINED.
DR EMBL; M68926; AAA41529.1; JOINED.
DR EMBL; M68927; AAA41529.1; JOINED.
DR PIR; A49744; A49744.
DR HSP; P22888; ILUT.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR000372; LRR_Nterm.
DR Pfam; PF00001; 7tm_1; 1.
DR Pfam; PF00560; LRR; 2.
DR PRINTS; PR00237; GPCRHHODPSN.
DR SMART; SM00013; LRRNT; 1.
DR PROSITE; PS00237; G_PROTEIN_RECP_FL_1; 1.
DR PROSITE; PS0262; G_PROTEIN_RECP_FL_2; 1.
DR G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
KW Phosphorylation; Repeat; Leucine-rich repeat; Alternative splicing.
FT SIGNAL 1 26
FT CHAIN 27 700 LUTROPIN-CHORIOGNADOTROPIC HORMONE
FT DOMAIN 27 362 RECEPTOR.
FT TRANSMEM 363 390 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 391 399 1 (POTENTIAL).
FT TRANSMEM 400 422 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 423 443 2 (POTENTIAL).
FT TRANSMEM 444 466 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 467 486 3 (POTENTIAL).
FT TRANSMEM 487 509 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 487 509 4 (POTENTIAL).

Query Match 16.2%; Score 134; DB 1; Length 700;
Best Local Similarity 27.2%; Pred. No. 5.9e-05;
Matches 34; Conservative 31; Mismatches 48; Indels 12; Gaps 4;

QY 1 AQYISVAIFGINLAFTIIIVFSGSMFYSVHOSAITATEIRNOVKEMILAKRFFIIVF 60
Db 528 SQVILSILI-LNVAVFVICACYIRIYFVQNPETAP-----NKDTKIAKKMAILIF 580
QY 61 TDALCWIPF-FVAKPLSLQLVEPGTITSWVIGSAINSNALNPILYTLTRFKE---- 115

```

Db 581 TDFTCADISFAISAKVPLITVNSKILLVLTPVNSCANPLYAIFTKARQDFELL 640

Qy 116 MHFR 120

Db 641 LLSRF 645

RESULT 15

FSHR\_MOUSE STANDARD; PRT; 692 AA.

AC P35378; Q904C2; Q9QW8;

DT 01-JUN-1994 (Rel. 29, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Follicle stimulating hormone receptor precursor (FSH-R) (Folliotropin receptor).

GN FSHR.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxId=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=129/Sv; TISSUE=Testis;

RA Tena-Sempere M., Manna P.R., Huhtaniemi I.T.;

RT "Molecular cloning of the mouse follicle stimulating hormone receptor complementary deoxyribonucleic acid: functional expression of alternatively spliced variants and receptor inactivation by a C566T substitution in exon 7 of the coding sequence."

RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Testis;

RA MEDLINE=21085660; PubMed=11217851;

RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y., Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S., Arizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamakura I., Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R., Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T., Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H., Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J., Schirni L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T., Sakai K., Okido T., Furuno M., Aono H., Baidarelli R., Barsh G., Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F., Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M., Gustincich S., Hall D., Hofmann M., Hume D.A., Kamiya M., Lee N.H., Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P., Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N., Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F., Suzuki H., Torio-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L., Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S., Hayashizaki Y.;

RA "Functional annotation of a full-length mouse cDNA collection."

RL Nature 409:685-690(2001).

RN [3]

RP SEQUENCE OF 1-51 FROM N.A.

RC MEDLINE=93093308; PubMed=1459341;

RA Huhtaniemi I.T., Eskola V., Pakarinen P., Matikainen T., Sprengel R.;

RT "The murine luteinizing hormone and follicle-stimulating hormone receptor genes: transcription initiation sites, putative promoter sequences and promoter activity."

RL Mol. Cell. Endocrinol. 88:55-66(1992).

CC -!- FUNCTION: RECEPTOR FOR FOLLICLE STIMULATING HORMONE. THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE ADENYLATE CYCLASE.

CC -!- SUBCELLULAR LOCATION: Integral membrane protein.

CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS. FSH/LSH/TSH SUBFAMILY.

CC -!- SIMILARITY: Contains 6 leucine-rich (LRR) repeats.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; AF095642; AAC67559.1; -

DR EMBL; AK016635; BAB30351.1; -

DR EMBL; S49632; AAB24401.1; -

DR EMBL; M87570; AAA37541.1; -

DR MGD; MGI:95583; Fshr.

DR InterPro; IPR000276; GPCR\_Rhodops.

DR InterPro; IPR001611; LRR\_Nterm.

DR InterPro; IPR000372; LRR\_Nterm.

DR Pfam; PF00001; 7tm1.1; 1.

DR Pfam; PF00560; LRR; 4.

DR Pfam; PF01462; LRRNT; 1.

DR PRINTS; PR00237; GPCRRHODOPS.

DR SMART; SM00013; LRRNT; 1.

DR PROSITE; PS00237; G\_PROTEIN\_RECF\_FL1; 1.

DR PROSITE; PS00262; G\_PROTEIN\_RECF\_FL2; 1.

DR G-protein coupled receptor; transmembrane; Glycoprotein; Signal;

KW Phosphorylation; Repeat; Leucine-rich repeat.

KW SIGNAL

FT CHAIN 1 17 POTENTIAL.

FT DOMAIN 18 692 FOLLICLE STIMULATING HORMONE RECEPTOR.

FT DOMAIN 18 365 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 366 386 1 (POTENTIAL).

FT DOMAIN 387 397 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 398 420 2 (POTENTIAL).

FT DOMAIN 421 442 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 443 464 3 (POTENTIAL).

FT DOMAIN 465 484 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 485 507 4 (POTENTIAL).

FT DOMAIN 508 527 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 528 549 5 (POTENTIAL).

FT DOMAIN 550 572 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 573 596 6 (POTENTIAL).

FT DOMAIN 597 607 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 608 629 7 (POTENTIAL).

FT DOMAIN 630 692 CYTOPLASMIC (POTENTIAL).

FT REPEAT 44 68 LRR 1.

FT REPEAT 69 93 LRR 2.

FT REPEAT 119 143 LRR 3.

FT REPEAT 170 192 LRR 4.

FT REPEAT 193 216 LRR 5.

FT REPEAT 218 240 LRR 6.

FT DISULFID 441 516 BY SIMILARITY.

FT CARBOHYD 191 191 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 199 199 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 293 293 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CONFLICT 436 436 Q -> K (IN REF. 2).

SC SEQUENCE 692 AA; 77769 MW; 4B57229180563844 CRC64;

Query Match 16.0%; Score 132.5; DB 1; Length 692;

Best Local Similarity 26.6%; Pred. No. 8e-05;

Matches 42; Conservative 30; Mismatches 57; Indels 29; Gaps 5;

Qy 1 AQIYSVAIFGINLAFFIIVFSGMFYSVHQSATATEIRNQVKEMILAKREFEIVF 60

Db 526 SOLYVMAL-LVLNALAFVVGCGTYHILVRNPNVSS-----SRDTKIAKRMATLIF 578

Qy 61 TDALCWPII-FVAKPLSLQLQVEIPGTITSMWVIGYSAINSAINPILYTLATRPFKF--- 115

Db 579 TDFLCMAPIUFAISLAKVLIVTSKAKILLVLFYPINSCANFLYAITKFRDFEV 638

Qy 116 MHFR-----WNYRQRKSMDSKIR 136

Db 639 LMSVFGCYEQAIYKTSTSTTHFNHRKNPCSSAPR 676

Search completed: September 25, 2003, 14:59:53

Job time : 12 secs